

**IBN HALDUN UNIVERSITY  
SCHOOL OF GRADUATE STUDIES  
DEPARTMENT OF ECONOMICS**

**MASTER THESIS**

**TURKEY LABOUR MARKET:  
REAL-TIME DATA COLLECTION AND INSIGHTS**

**VAHID FARAJIJOBEHDAR**

**THESIS SUPERVISOR  
ASST. PROF. ASAD UL ISLAM KHAN**

**ISTANBUL, 2022**

**IBN HALDUN UNIVERSITY  
SCHOOL OF GRADUATE STUDIES  
DEPARTMENT OF ECONOMICS**

**MASTER THESIS**

**TURKEY LABOUR MARKET:  
REAL-TIME DATA COLLECTION AND INSIGHTS**

**by**

**VAHID FARAJIJOBEHDAR**

**A thesis submitted to the School of Graduate Studies in partial  
fulfillment of the requirements for the degree of  
Master of Arts in Economics**

**THESIS SUPERVISOR**

**ASST. PROF. ASAD UL ISLAM KHAN**

**ISTANBUL, 2022**

APPROVAL PAGE

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Arts Economics

Thesis Jury Members

Title - Name Surname

Opinion

Signature

_____	_____	_____
_____	_____	_____
_____	_____	_____

This is to confirm that this thesis complies with all the standards set by the School of Graduate Studies of Ibn Haldun University.

Date of Submission

Seal/Signature

## ACADEMIC HONESTY ATTESTATION

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Surname:

Signatur



ÖZ

TÜRKİYE İŞ PİYASASI: GERÇEK ZAMAN (BİG) DATA YÖNTEMİ İLE  
VERİ TOPLAMA VE ÖNGÖRÜLER

Farajjobehdar, Vahid

İktisat Yüksek Lisans Programı

Öğrenci Numarası: 194024011

Open Researcher and Contributor ID (ORC-ID): 0000-0002-9153-0765

Ulusal Tez Merkezi Referans Numarası: 10505635

Tez Danışmanı: Asst. Dr. Asad Ul Islam Khan

Dec 2022, 52 sayfa

İşgücü piyasasının dinamikleri sık sık değişilmesine rağmen, veri toplama yöntemleri henüz geleneksel metodoloji ile yapılmaktadır. Bu geleneksel anket analizlerin yeterli olmadığı nedeniyle ücret değişimi, işgücü arz ve talep veya eğitim ihtiyaçlar gibi faktörler doğrudan veriye ulaşma konusunda büyük hatalar içerebilir. Bu nedenle veriye dayalı karar çağında olduğumuz devirde artık teknoloji dönüşüm ile çözüm bulmalıyız. Bu tezin temel hedefi, Türkiye'deki reel işgücü piyasasını araştırarak, büyük veri toplama metodolojisi, veri kalıpları ve mimari önermeyi amaçlamaktadır. 2021-22 haftalık düzenli veri toplama sürdürerek, 314 bin iş ilanı incelendi ve maaş açıklığı, yetenek ve piyasa öngörülerini araştırıldı. Sonuçlar, Türkiye'nin işgücü piyasasının önümüzdeki yıllarda özellikle covid-19 etkisi ile, sektörde ve iş pozisyon arasında, yeni becerileriyle önem verilmesine, piyasa verilerin öngörü ve analizini kısa vade olarak bir standard uygulanmasını ve özel istatistik öngörü ile analizlerini önermektedir.

**Anahtar Kelimeler:** Big Data, İstihdam, İş Piyasası, İşgücü Verileri, Yetenğe Dayalı-Ekonomi.

## ABSTRACT

### TURKEY LABOUR MARKET: REAL DATA COLLECTION AND INSIGHTS

Farajijobehdar, Vahid

MA in Economics

Student ID: 194024011

Open Researcher and Contributor ID (ORCID): 00000-0002-9153-0765

National Thesis Centre Reference Number: 10505635

Thesis/Project Supervisor: Asst. Prof. Dr Asad Ul Islam Khan

Dec 2022, 52 Pages

Still, there is a conventional approach for collecting data on the dynamics of the labour market, such as salary variation, supply and demand, mobility, or insights into educational needs and knowledge workers. Hence, more than the traditional survey analysis method is needed and can mislead the reality as we are in an era of exponentially increasing data.

This thesis aims to propose a data collection methodology in real-time, data patterns, and architecture to deal with day-to-day labour market variation in Turkey. During one year of regular data collection, more than 314 thousand job titles in 2760 occupations are analysed regarding the pay gap, education, and preferences. The results show that Turkey's labour market is shifting, especially after the Covid-19 effect on the skilled-based economy, digital transformation, and payroll distribution index in the upcoming years.

**Keywords:** Big-Data Collection, Employability, Labour Market Data, Skill-Based Economy.

## DEDICATION

This thesis is dedicated to people who need help finding decent jobs and have a barrier to educating and upskilling themselves. This a big wish, but we hope that in the upcoming years, the trend line of demand and supply will be in a good equilibrium to have near-zero unemployment, illiteracy, and a lack of resources. This wish sounds emotional, but I hope we can experience rational growth planning to target the overall welfare.

This thesis literally cannot change the world of employment. Nevertheless, even having an idea from a different perspective deserves time and work to research. I am grateful to have the privilege of studying in Turkey and working in the recruitment industry. I also dedicate the thesis to my professors, colleagues, and friends in Turkey.

All best regards.

## ACKNOWLEDGEMENT

It is a privilege to have the sincere support of Ibn Haldun University academia, the economics department, and resources in Türkiye from higher education terms. I want to thank Ibn Haldun University and the community, my professor Asad Ul Islam Khan, the Graduate Office, the juries of the thesis defence, the Dean of our department, classmates and my parents and wife (Bahar) for sharing the time and hand to write, think and finish this master's program.

I appreciate Nima and Ahmet, my colleagues, who have encouraged and facilitated this study by recommending technologies to code and collect data.

Last but not least, Kariyer.net and my colleagues there should be praised as they are anchoring data to collect and extract insights which, with their work, this study is meaningful.

I hope not to forget other names, but I want to thank everyone who has contributed to this modest academic research.

Vahid FARAJIJOBEHDAR  
ISTANBUL, 2022

## TABLE OF CONTENTS

<b>ÖZ</b> .....	<b>iv</b>
<b>ABSTRACT</b> .....	<b>v</b>
<b>DEDICATION</b> .....	<b>vi</b>
<b>ACKNOWLEDGEMENT</b> .....	<b>vii</b>
<b>TABLE OF CONTENTS</b> .....	<b>viii</b>
<b>LIST OF TABLES</b> .....	<b>x</b>
<b>LIST OF FIGURES</b> .....	<b>xi</b>
<b>LIST OF ILLUSTRATIONS</b> .....	<b>xii</b>
<b>LIST OF SYMBOLS AND ABBREVIATIONS</b> .....	<b>xiii</b>
<b>CHAPTER I INTRODUCTION</b> .....	<b>1</b>
1.1. Trends in Labour Market Data .....	1
1.2. Turkey Labour Market Perspective .....	2
1.3. Scope of the Study.....	3
<b>CHAPTER II LITERATURE REVIEW</b> .....	<b>4</b>
2.1. Turkey's Labour Market in Brief.....	4
2.2. Antecedent Studies: Labour Market Analysis in Turkey .....	5
2.3. Conditions in Job Market Equilibrium.....	7
2.3.1. Job Preference .....	7
2.3.2. Pandemic Effect and the Transition from Time to Delivery .....	8
2.3.3. The Queue Analogy in Job Preference .....	9
2.3.4. Salary: Always as an Important Factor .....	9
2.3.5. Education, Training vs Skill-data in a Labour Economy .....	10
2.4. Limitations of Traditional Methodology for Collecting LMI Data.....	11
2.5. The Online Platforms in the Job Market: Why We Select Data from the Platform of Kariyer.net.....	11
2.6. The Application of Big-data Collection for LMI in Turkey .....	12

2.6.1. Relational Modelling from LMI in the Turkey Market.....	12
2.6.2. Deriving Insights into Turkey's Labour Market .....	14
<b>CHAPTER III METHODOLOGY AND APPLICATION.....</b>	<b>16</b>
3.1. Methodology of Research.....	16
3.2. Collection Procedure .....	16
3.3. Data Patterns (Relational Model) .....	18
3.4. Data Analysis .....	20
3.5. Technical Tools .....	22
3.5.1. Web Scraping by Python Scrape Library .....	22
3.5.2. Relational Data Modelling by PostgreSQL.....	23
3.5.3. Extracting Insights and Visualisation.....	24
<b>CHAPTER IV RESULTS AND DISCUSSION .....</b>	<b>25</b>
4.1. General Insight .....	25
4.2. Preference of Job Market (Supply Side) .....	26
4.3. Employment Status (Demand Side) .....	28
<b>CHAPTER V CONCLUSIONS .....</b>	<b>33</b>
5.1. Conclusion.....	33
5.2. Practical Recommendations .....	33
5.3. Suggestions for Future Studies .....	34
5.4. Limitations of Thesis.....	35
<b>REFERENCES.....</b>	<b>36</b>
<b>APPENDIXES .....</b>	<b>39</b>
<b>APENDIX A.....</b>	<b>39</b>
<b>APENDIX B.....</b>	<b>49</b>
<b>CURRICULUM VITAE.....</b>	<b>51</b>

## LIST OF TABLES

Table 2.1. Selected Studies on Labour Market and Effective Parameters .....	6
Table 2.2. Relational Model of Data for Collection and Extraction Insights .....	13
Table 2.3. Data Structure for Scraping the Data of Kariyer.net.....	14
Table 3.1. Type of Data and Properties in Kariyer.net for Occupations.....	19
Table 3.2. Type of Data for Relational Modeling.....	20
Table 4.1. Aggregate Data for Employment Demand.....	27
Table 4.2. Number of Undergraduate Students Registered for 2017-18, Turkey .....	30
Table 4.3. Number of Graduate Students Registered for 2020 -21, Turkey .....	30
Table 4.4. Number of Vocational Training by Iskur, Jan-Dec 2021.....	31
Table B.1. Industry Ranking Based on Collected Data .....	49
Table B.2. Education Status Rankings.....	50
Table B.3. Demanding Skills Rankings .....	50

## LIST OF FIGURES

Figure 2.1. Unemployment Trends (Percentage*) in Turkey, 2020-22 .....	5
Figure 2.2. Employment Status (Scale of Million*) in Turkey, 2020-22 .....	5
Figure 3.1. Relational Model for Data Analysis of Labour Market Information.....	21
Figure 4.1. Relational Data of Scraped Tables .....	25
Figure 4.2. Employment Demanding (Open Positions to Apply).....	27
Figure 4.3. Job Titles Frequency on Unique Titles in Different Sectors .....	29
Figure 4.4. Education and Training Needs of Labour Market .....	29
Figure 4.5. Skill Frequency Across Positions .....	32



## LIST OF ILLUSTRATIONS

Illustration 3.1. Kariyer.net Position Listed Page in Alphabetic Order .....	17
Illustration 3.2. Scrapyng Data Models and Step Codes by VS Console .....	18
Illustration 3.3. Connection to PostGre to Collect Data .....	23
Illustration 3.4. PgAdmin Dashboard for Kariyer.net.....	24
Illustration 3.5. Relational Data of Scraped Tables .....	24



## LIST OF SYMBOLS AND ABBREVIATIONS

CEDEFOP	European Centre for Development and Vocational Training
CSS:	Cascading Style Sheets
DESI	Digital Economy and Society Index
ESCO	European Skills, Competences, Qualifications and Occupations
İŞKUR	Türkiye İş Kurumu (Turkey Bureau of Labour Force)
GÖÇ	Turkey immigration office
HTML	Hypertext Markup Language
ILO	International Labour Organisation
LMI	Labour Market Information
NESTA	UK National Endowment for Science, Tech and Arts
SQL	Structured Query Language
TÜİK	Türkiye İstatistik Kurumu (Turkey Bureau of Statistics)
WEF	World Economic Forum
YÖK	Turkey Higher Education Organization
$t_n$	Time
$\dot{E}_n$	Expertise
$\dot{W}_n$	Wage

# CHAPTER I

## INTRODUCTION

### 1.1. Trends in Labour Market Data

Traditional surveys, data banks, yearly calendars, and published labour market statistics by private or state organisations can quickly be doomed as employment dynamics change swiftly. The nature of work and employability as a batch of data can be collected on a large scale; however, the parameters of employment, workforce characteristics, salary projections, and related data are much more challenging to reach in a synchronised way. For instance, coming to required skills or competencies, there is no universal language to define which skills are most required in similar jobs to plan for education and meet industry needs. On the other hand, estimating salaries involves many outputs that require deep analysis and data collection. These issues can hinder the fundamental analysis of the labour market. In addition, all related data to the labour force (like employment rate, wages, qualification, or education-degree status) still are collected in a very conventional way, therefore, are disconnected from the real parameters of labour data.

From a different perspective, a skill-based economy is the near future topic of labour market data. According to the LinkedIn economic graph (2022), the nature of the job market is changing gradually, flexibility and benefits are topics to spot, and the skill for jobs will vary by up to 40% by 2027. There is an estimation that digital technologies might transform more than 1 billion jobs worldwide; in simple words, 25% of skills are doomed or less required in employment. It means that skill employment would be the core part of the labour market more than ever. World Economic Forum from 2019 up to the upcoming year (2023) devotes to closing the skill gap to make people aware of future trends in the skill economy.

In the labour market, concepts such as green skills, talent mobility, shortage or surplus are familiar, but the approach and perspective of data analysis are quite different from the 2000s. Digitalisation, transformative and massive IT automation is essential in future research studies about the labour market. Some specific organisations like NESTA in the UK routinely study this kind of approach to boost labour market intelligence. This institution explicitly calls for market intelligence (Nesta, 2020) to provide indices and metrics about the local market of the UK. Although many researchers and institutions still prefer conventional methods to extract the report, data, and insight. Coming to Turkey, in this thesis, we aim to bridge the gap from a conceptual study of data patterns and methodology of real data and bring real insight to complete the big picture of how the labour market experiences volatility.

In addition, necessary data like salaries, open job vacancies, or adverts showing the employment trend in the labour market, qualifications, and later skills are rarely collected historically by conventional surveys, direct interviews, or any typical data collection. Moreover, one of the survey's problems and limitations is the participation rate (Bethlehem J. et al. (2021)) which leads to a direct or indirect negative impact on results. Regardless of the accuracy, collecting data in a 1:1 format takes time and effort. Data latency and accuracy are very important in a turbulent and volatile market like Turkey. Crowdsourcing data collection is trending in an engaging format, like smart check-ins and short feedback. The comments in e-commerce platforms like Hepsiburada, Trendyol or any rate pulse question on large scales work better than a one-time collecting data procedure. The continuity and scalability of the data collection make it possible to validate the data from different sources. These thoughts are our main concerns and, of course, motivation to investigate the labour market data of Turkey not only through the lens of industry, occupation and salaries but also from the education and required skills and qualifications across the advertised jobs on career platforms and other live data sources.

## **1.2. Turkey Labour Market Perspective**

As the inflation rate, exchange rate, supply, demand or other exogenous or endogenous variables are reflected in the labour market parameters like employment, salaries, and mobility, in this study, we aim to collect, clean, and analyse the time-interval data

rather than ready-to-use surveys. To understand the possibility of larger data collection and extract real insights into the labour market in Turkey, we choose the large scale of data which can be collected automatically by technology. Instead of using the old data, we try to minimise the time of data collection and initial analysis. This procedure can transform meaningful insight from multiple resources to understand the real puzzles of creating jobs (geographically), predicting mobility, and gauging demanding or dying skills to plan better for near employment or even education adjustment. Overall, the research mostly emphasises the methodology and pattern of data rather than complex statistical inferences, as the survey analysis has limitations in line with future trends.

### **1.3. Scope of the Study**

In the literature review, we will discuss the main data pattern of labour market information (LMI) with a systematic approach. Then we try to propose relational data models in Turkey's labour market. The parameters rarely discussed are job application preferences, geographical mobility, demand, surpass of skills, and what area can create an imbalanced market. The extension of the data collection model using technologies like Structural Language Query (SQL), Web data extraction (web scraping), and munging data with R (aggregating data of salaries, skills, occupation, and roles etc.) are the main points which will be addressed to.

Chapter 3 covers how to select the type of data, new survey models like pulse or check-ins, the aggregate of salaries and occupation data, and descriptive statistics outlined in 2 parts. In Chapter 4, we identify the main insight of collected data and salaries, the shifts, differences across industries, pay gap across positions and geographical indices and try to conclude in Chapter 5 to wrap all these plus future suggestions and limitations of the study. All codes, important data models and SQL queries are listed in the Appendix, ranging from lab.py codes to relational database models.

## CHAPTER II

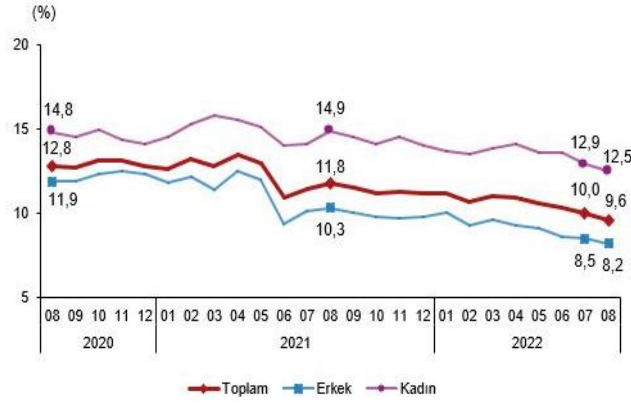
### LITERATURE REVIEW

#### 2.1. Turkey's Labour Market in Brief

Turkey is the eighteenth most populated country in the world, where nearly 85 million people live here (Tüik, 2021), with the highest infantile rate (26%) and youth population (15,3%) among all European Union countries. Turkey is the first young country compared to all European countries. Although the population age and birth rate are hot topics of Turkey's future, Türkiye can still provide the labour needed in this region. The birth rate is 1.7%, ranking seventh among the European Union. Turkey has been a proxy of mass immigration since 2011 and is known as one of the most receiving countries of immigrants and refugees more than 3.6 million refugees plus 1.7 million issued residence permits (Göç, 2021; Mülteciler derneği, 2022). To these numbers, it can be added unregistered people where this is no data. Undoubtedly, there are risks and opportunities through immigration, but we excluded the immigration effect as we have not too focused on and provided the contextual data.

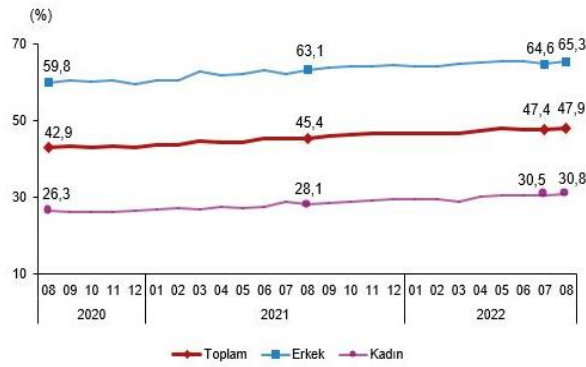
The unemployment rate of Turkey is discussible in terms of multiple conditions. The overall unemployment, according to the Turkey Bureau of Statistics (Tüik, October 2022, last accessible release), has experienced a downward trend after Q3 2021 (modified by the pandemic effect and layoffs). The seasonality effect can be considered one of the parameters in overall employment/unemployment. Tüik data (2022) has projected a positive trend for decreasing unemployment this year. Nevertheless, the main questionable variable is labour complexity and job preferences.

This chapter will discuss how the labour market is affected by preference, the creation of jobs, the demand and supply sides of the market, and technological advances. Figures below shows the unemployment and employment rates in recent two years:



**Figure 2.1. Unemployment Trends (Percentage\*) in Turkey, 2020-22**

Source:TÜİK (2022)



**Figure 2.2. Employment Status (Scale of Million\*) in Turkey, 2020-22**

Source:TÜİK (2022)

## 2.2. Antecedent Studies: Labour Market Analysis in Turkey

Labour market analysis in Turkey has been the subject of many studies, locally or internationally. Although many recent topics discussed immigration and refugees' factor as influencing factors in the labour market of Turkey (Loyaza et al., 2018; Aksu et al., 2022; Del Carpio, 2017), the local researcher proposed models and estimation for labour market dynamics. Akçomak, S., & Gürçihan, H. B. (2013) have studied the education factor and its importance in the service industry. Turkish scholars mainly discussed wage discrimination (Cudeville, E., & Gurbuzer, L. Y. (2010)) mass immigration and the impact of Syrian refugees on the labour market (Ahmet T. Akcan, 2018; Ege Aksu et al., 2022) and growth and economic performance (Akçoraoğlu, A.

(2010)). Gündüz Höşgör and Jeroen Smith (2008) discussed how gender participation could influence the labour market from multiple perspectives. The findings show that the effect of modernisation on female workers varies by supporting their hypotheses. Apart from the socio-economic studies, other scholars, Mahmut Özer et al. (2020) and İşık V. (2016) paid more attention to the education and labour gap. The results confirm the role of skill mismatch and education on the employability index.

The wage gap and population are other important factors to consider in this dynamic economic theme of Turkey (Aslan G., 2016). Ipek Ilkaracan and Raziye Selim (2007) found that the source of the wage gap, considering gender inequality, comes from levels of experience and the working context. In line with Turkish labour market studies, experience, skills, and wages or salaries are the most concerned topics when we review the last ten years of research. Data play a detrimental role in analysing and projecting all interrelated factors at this scale. We are in the age of data (Einav, L., & Levin, J. (2014)), where the multiple sides of variables can be thoroughly investigated. However, big data is not a panacea to heal all the challenges of data gathering and afterwards. Data mining or harvesting is a very important factor, and then after, the cleaning part. In addition to Turkish scholars, the below table briefs the main related topics of the labour market from other scholars.

**Table 2.1. Selected Studies on Labour Market and Effective Parameters**

Researcher(s)	Year	Topic of study	Results
Deniz Sevinc, Anne Green, John R. Bryson, Simon Collison, Rebeca Riley and Simon Adderley	2020	Skills and effect on regional development in the UK	Skills are regarded as the key driver of industry boost at the local level. It is recommended that a methodology to estimate the qualification alignment with the demand of supply in the labour market
Wassily Leontief	1986	Input and Output economy (Wages, Profits, prices and taxes)	It gives some good examples to show the direct and indirect dependency of productivity and price index on wages, which is very complex to estimate. Price changes across industries or production lead to a wage rise, but the relationship could be more direct.

**Table 2.1. (cont.)**

Imane Khaouja, Ghita Mezzour, Kathleen M. Carley, Ismail Kassou	2019	Recommending a skill taxonomy adopted from job openings in French and English languages	The focus of the research is to propose the automation process to identify skills in a large set of data by the SPARQL endpoint. The results show increased accuracy in identifying top soft skills for universities to train students' competencies.
Tulika Bhattacharyaa, Bornali Bhandari, Indrajit Bairagyac	2020	Estimating skill-based employment across industries in India.	The paper defines skill employment in 4 different categories ranging from low to high skilled. The assumption is a fixed coefficient to different types of skills to predict employment linkages. The findings unveil that skill composition will change over time as passing the time affects the demand and supply in the market.

### **2.3. Conditions in Job Market Equilibrium:**

#### **2.3.1. Job Preference**

State and private sectors research labour market periodical analysis in Turkey from an economic perspective. Regarding detailed and real-time analysis, most studies aim to determine the gap in labour market parameters like wage, demand, or supply or wholesale of labour market ingredients (*Johnson, E. (2016)*). However, it is time to scrutinise the detrimental factors that create volatility, such as job preference, mobility and geographical effect on creating employment. In this assumption: job preference is an intention to apply moderated by behavioural factors. The preference can be defined from multiple states: a) active job seekers (not being employed nor having any past employer) like fresh grads or students or any individuals without work experience 2) individuals who want to upscale their career. In this model, the deal between the job seeker and the employer is enacted when the overall condition is satisfied in the simple below equation (*Bhattacharya et al., 2020*):

- Time (to serve)

- Expertise (to solve the problem in a better way)

In a simple equation  $\dot{t}_n + \dot{E}_n = \dot{W}_n$

As salary in the market is calculated by dedicated time (hourly, daily or monthly) according to law, the overall perception is the exchange of time to fill the tasks. Now with the emerging transformation, the equation can be more complex. The flexibility of time can be discussed in the following parts; however, some employers now expect delivery than fill the time. We only generalise it to some sectors like production, but the pandemic broke the sanity of being only at the office. In return, job preference to apply for employment differs from before the pandemic.

### **2.3.2. Pandemic Effect and the Transition from Time to Delivery**

Now in our classification, there are three main groups of job seekers: a) new grads from any level of degree, b) employees who are already working but want to move up in the same career path, c) individuals who want to change careers from a different field or background in internal d) c condition as external mobility (lay off effect). This classification shows that preference in career selection is much broader. Soon after the pandemic, working from home and remotely were added to HR jargon, and people felt they could work alone, leading to big resignations. Less than one year, the layoffs have started as the effect of over-hire in some sectors. LinkedIn economic graph (2022) illustrates that the condition is better, unlike before. Quantifying job preference is very complex. The only parameter we can understand is the quantification of application per open or advert job on any recruitment channel. Across industries, we can understand the factors influencing people's preferences with different tests and surveys. However, in this thesis, the only quantitative way to see people's preferences is the application number per job position. Thus, we should consider people's preferences (job seekers in any state-how much they would like to apply to open job adverts) to see the demand side from the employee dimension.

Job preference in the market varies as a job seeker's behaviour varies due to financial status, expectation, background or even self-efficacy or confidence in doing jobs (Tansel, A. and Acar, E.Ö. (2017)). In this approach, we only consider the preference as a tendency to apply without knowing any behavioural reasons. This is one of the most important aspects of real-time data collection by which we can understand the

more dynamic side of the labour market. Generally, the number of applications is disregarded when unemployment or employment is analysed.

### **2.3.3. The Queue Analogy in Job Preference**

Having noted that human desire tends to reach a higher hierarchy, it is normal to hear people want to take higher positions. This is a simplified analogy that when we wait in a bus station, depending on our time scarcity, we can change our expectation to take a seat or stand to arrive on time. Sometimes, people prefer to get on a less crowded bus to take a seat. Passing the time or scarcity of available buses may affect people's decision to take a bus. It can be the same in the labour market: people may disregard some positions due to different expectations, but if they know that there is a limitation in the job market or a high possibility of layoffs occurring in defined positions, they reconsider the expectation. We aim to see how many positions are less considered during the past year and if sufficient job availability exists. Time is an important parameter in the labour market which can be highly trackable in real-time analysis.

### **2.3.4. Salary: Always as an Important Factor**

Regarding salary calculation in the market, many Turkish economists consider minimum wage (asgari ücret). Asgari in the Arabic-root lexicon means minimum, and interestingly we do not hear the Azami (maximum) wage frequently. The state adjusts the minimum wage annually considering the inflation rate (15), but the foreign exchange rate and other macroeconomic factors are inevitable. Salaries can be divided into two ways of calculation: Standard/conventional level compensation according to the Organization of Labour for state jobs like military, higher education and health. Secondly, the private sector struggles to define a specific or one-size for all compensation system. Putting these two pieces on one side shows that job classifications cannot bear the scalability and generalisation of salary adjustment.

Platforms like Glassdoor (US) and PayScale provide salary projections so people can know whether to accept or reject offers. In Turkey, salary is a very sensitive topic. It is very personal, and normally people hesitate to ask as it is considered confidential. Nevertheless, how do people, at least in Turkey, determine the market value and propose expectations? In chapter 4 and after the insight of Turkey, we will discuss this

question from multiple perspectives. This is the second concern that this thesis tries to shed light on it. Whether there is meaningful data or not, how should it be addressed as the main parameter of the labour market?

### **2.3.5. Education, Training vs Skill-data in a Labour Economy**

The current president of the European Union, Ursula von der Leyen, is trying to nominate the upcoming year of 2023 as the "European Year of Skills". The report of The Digital Economy and Society Index (DESI) by the European Commission highlights those digital skills are an indispensable chain of future work in the connected world and skill-based labour market analysis ignited by the European Centre for Development and Vocational Training. The shortage of the required market is one of the main concerns they want to deal with. According to Tulika Bhattacharya et al. (2020), the nature of skills is changeable across time variability as the demand for skills is not defined directly by educational status. The level of skills and need for people is mostly defined by the nature of the industry. This shows how we need skill-based economy data to estimate the whole labour market picture. Skills are not bound to education.

In contrast, any competency to deliver tasks and projects and handle the job at a mastery level can be labelled as skills. More than ever, we hear the skill tech platform to understand what skills and what types of skills exist in the market to match the talent or labour force to projected jobs. Giant tech companies like Microsoft (2022) have invested in skilled tech platforms which sooner or later affect the middle east conception of the labour market. The most interesting part is policy and economic research like UK, EU and US. In this regard, the European Union launched a standard taxonomy called ESCO (2017). Economic modelling in the United States and JobKred are other initiatives to boost economic growth and employment in the mentioned regions. Here the skills, qualifications and demand side are important, and all goes on negotiation style. People search on platforms to define and feel what they can deserve.

## **2.4. Limitations of Traditional Methodology for Collecting LMI Data**

When we look at Turkey's Labour market information (LMI), we can find only employment rate, education status and salaries. Given International labour organisations (Maas et al., 2018), more than the traditional surveys are required to reveal all faces of labour market information. Using the big data approach to supplement the labour market is a way to break down the puzzles of employment and demographic dynamics. Time is gone to only refer to job titles as the normalisation and skill-powered economy is highly appreciated.

The employment policy is adjustable quickly to invest in education, planning for universities, high schools, vocational training or even state-based employment. If data is usable and understandable, we can have a good mix methodology of collecting the right data and, therefore, a depth analysis. Even though it is not straightforward to analyse the job descriptions, matching the syllabus with industry or any forecasting, thanks to the algorithm of natural process language \*(NPL) and machine learning techniques, now are not considered futuristic trends. The limitations of traditional survey methodology pushed us to look into the labour market information in Turkey from a big data perspective. In the following parts, we will discuss the main limitations of this traditional collecting information and the opportunities of big data as a supplementary in-depth analysis of Turkey's job marketplace.

## **2.5. The Online Platforms in the Job Market: Why We Select Data from the Platform of Kariyer.net**

There are multiple channels for online recruitment, job boards or announcements, CV makers or career platforms which try to facilitate employment and recruitment in Turkey. To mention a few, LinkedIn, Kariyer.net, İşin Olsun, Glassdoor, indeed (a search engine for posted jobs), Yetenek Kapısı (by the human resource office of the Turkish government), Secret Cv, Toptalent etc. There are multi-players in the employment or recruitment ecosystem in Turkey, but the market size is in the hand of kariyer.net, with 25 million CVs and monthly 32 million visits. iLab ventured start-up, Kariyer.net is one of the leading HRtech in Turkey and holds 10m+ app downloads

(data claimed by iLab, 2022) on two different segments of blue and white-collar job seekers/employees.

According to Kariyer.net, which collaborated with PAL (2022), the employable population increased to 64 million in 2021. The total size of job orientation (from application to landing a job) has been reported to be 28% considering two platforms of Kariyer and İşin Olsun. If this number is reliable, then one-third of employment occurs by facilitation of kariyer.net. 13.8 million in the private sector out of 26,8 million employed individuals at the end of 2020. 2.6 million job seekers in Turkey found their job through kariyer.net annually during 2019-21. This is the main reason to choose Kariyer.net as it holds more users, job announcements and comprehensive labour statistics like salaries, education and preferred skills compared to other rivals.

The online job board and recruitment funnel has another application called İşin Olsun for blue-collar workers. The reason for choosing kariyer.net is the enrichment and structure of accessible data, e.g., salaries report, continuous job posting and daily updates, basic labour statistics and numbers like salaries, education status, capabilities and skills, which are the most valuable data to measure the dynamics of the labour market. However, it is inevitable to miss and lose the real job posts created by paper or published in conventional mediums like newspapers, TV etc. In this research, we target to track and collect the posted jobs published by Kariyer.net weekly. We started to understand the logic and structure of websites to collect data in real-time to minimise human errors and maximise the accessibility of data.

## **2.6. The Application of Big-data Collection for LMI in Turkey**

### **2.6.1. Relational Modelling from LMI in the Turkey Market**

If we try to connect the dots between surveys and real-time data of the LMI, the main requirement is creating relationships between the different data patterns. While we will highlight the data pattern and data collection method in Chapter 3, the logic and the relational model are elaborated across the body of research. By Structural Language Query, a relational model (technical schematic is given in the Appendix) is designed to collect data directly from the platform and analyse it afterwards.

- 1) **Occupations:** Kariyer.net lists 2670 occupations for grouping related job positions. We have a different label here than Kariyer (pozisyonlar), where the occupation leads the job positions or titles. In this regard, our data collection includes only those averted since 2021, when it started crawling. Only 1600 occupations hold reliable data related to salaries, education, skills, and position details. The relational model is designed given to these definitions:
- 2) **Job Positions:** weekly posted job opportunities through the website as a subset of occupation.
- 3) **Salaries:** the average, max and min salaries as people participated in job applications from Kariyer.net's survey.
- 4) **People data:** Number of people who participated in giving data of salaries and gender and educational levels.

**Table 2.2. Relational Model of Data for Collection and Extraction Insights**

Variable (Salary)	Salary average	Min salary	Max salary
Nature	Float Number	Float	Float
Variable (Preference)	Gender	Respondents per item	Occupations
Nature	Double. Per cent	Integer	Character

- **Skills:** the most required skills per occupation and position.
- **University:** Top universities in which people get hired for occupations.
- **Education and Vocational training:** the level of education of participating people in a job application. İŞKUR and YÖK data added too.
- **Major:** Listed majors that are employed

- **Sector:** The industry or area in which positions are in the market
- **Industry (department);** the department/industries for collected job positions.

**Table 2.3. Data Structure for Scraping the Data of Kariyer.net**

Variables (Skills)	Skill-occupation	University Name	Education status
Nature	Character, String	Character, String	String, Integer
Variables (Education)	Major	Sector-Industry	Department
Nature	Double. Percentage	Integer Number	

- Detail of positions (announced jobs): Roughly 2679 unique job positions in kariyer.net are scrapped weekly.
- Monthly variance: 10-15k (if the repetitive job adverts are excluded).

### 2.6.2. Deriving insights into Turkey's labour market

As discussed in this chapter, labour market information or factors constantly change due to the emergence or dying of required skills, knowledge, qualification, nature of work or technology, and demographic varieties. Last year we experienced a big resignation, and now we are witnessing the expected big layoffs, especially in the tech industry. More than 11 thousand in Meta (Facebook) is an example. Nobody should have predicted this, as digital transformation and investment in the metaverse were booming. This just happened less than one year! The Imbalance issue in the labour market leads to skill mismatch, unemployment, and surplus or shortage in demand and supply sides. If the market in Turkey is ready for future shifts in Tech or demographic changes, there will be a surplus in some industries and shortage in others.

The surveys have significant limitations in addressing these issues as they cannot track the changes in real-time. To define the indicators of the imbalance of the market, we need a robust big data pattern, a relational model, and more importantly, getting instant insights. With the emergence of open data conception, much iceberg-like information like demanding skills, needed education level, pay gap or other sets of this kind of data are automatically extracted from multiple online resources. The platforms of career planning, job announcements, higher education or national education published data, salary surveys, and projection of the labour market are the ecosystem of the LMI system. Extracting the imbalance rate is easier if we have all this data in one place. The analysis can be done in a very short period to derive the imbalanced labour market, and adjustment is applied immediately. This point is the researcher's main concern in proposing the real-time data collection model. Chapter 3 will discuss the data collection procedure of mentioned variables via a relational database model to create meaningful tables and extract insights accordingly.

## CHAPTER III

### METHODOLOY AND APPLICATION

#### 3.1. Methodology of Research

Collecting a large data set during a specific time is impossible with surveys or sample methods. Total labour participation is 53% which shows a huge growth compared to five years (PAL, 2021). In addition, on average, individuals in Turkey spend almost 7 hours on the internet, which hits 81% of internet penetration. (TÜİK, 2021). The basic research for any labour market analysis starts with strong data collection in the time series method. This part is one of the most important and, on the other hand, challenging parts of any good research. As discussed in detail in Chapter 2, the main contribution of this thesis is to provide the methodology for collecting data and insights from real data on the labour market. In the following lines, we structure the methodology and approach of the thesis, which is considered a novel way to catch more data resources in a short period, collect on a big scale by scraping, wrangling and cleaning data to make a relational model and lastly analyse it to extract insights. To collect data, we used web-scraping, then data wrangling and cleaning mainly and the relational model is designed by Structural Query Language (PostgreSQL).

#### 3.2. Collection Procedure

For data collection, first, we identified the main (URLs) and sitemap of Kariyer.net to understand the data pattern. Then the structure of the HTML and CSS of the website was thoroughly analysed to collect the data by spotting the right location of data. On the positions page, kariyer.net lists all occupations alphabetically, consisting of 2790 different position labels. The salaries, education and details of positions are completed by participation of applicants ranging from 50 to +999. Data collection started in March 2021 and was closed in mid-May 2022. To make the accuracy and clean the redundancy, the munge of data started in June and ended in September 2022. Every

week based on the calendar, we run the lib codes (fig 3.2) to collect details of data in the three main categories automatically:

- a) **Occupation**. All occupation parts were gradually scrapped and modified during the data collection procedure. In Python, we wrote class (step 1) and division designed to create the right path to design **structural data modelling**.
- b) **Occupation details**: different divisions of data such as salaries and participants
- c) **Position Details**: The date, company profile, location and frequencies of open jobs.

As illustrated below we analysed the position pages on Kariyer.net platform. There were more than 2700 indexed positions with details of salaries, skillset and education. Then, by power of Python and VS code library, we started to run codes from the selected pages.



The screenshot shows the Kariyer.net website interface. The browser address bar displays 'kariyer.net/pozisyonlar'. The website header includes the logo 'kariyer.net' and navigation links: 'İş Ara', 'Kariyer Rehberi', and 'Pozisyon Rehberi'. There are buttons for 'ÜYE GİRİŞİ' and 'İLAN YAYINLA'. The main content area features a banner with the text '2790 farklı pozisyon hakkında iş tanımından maaş bilgisine kadar tüm merak ettiklerini burada keşfet!' and a search bar with the placeholder 'Pozisyon Ara (Örn: Pazarlama Uzmanı)'. Below the search bar is a horizontal navigation bar with letters A through Z and 0-9. The main content area displays a list of job positions under the letter 'A':

Job Title
A Sınıfı İş Güvenliği Uzmanı
Abkant Operatörü
Acente Satış Uzmanı
Acil Tıp Teknisyeni
AB Proje Asistanı
Acente Müdürü
Acente Temsilcisi
Acil Tıp Uzmanı
AB Proje Geliştirme Uzmanı
Acente Müşteri Temsilcisi
Acil Servis Hemşiresi
Aday Mühendis

**Illustration 3. 1. Kariyer.net Position Listed Page in Alphabetic Order**

```

1 from scrapy.crawler import CrawlerProcess
2 from spider import OccupationDetailSpider, OccupationsSpider, PositionSpider
3 from context import Context
4
5 def step3_positions_spider():
6     urls = []
7     base_url = 'https://www.kariyer.net'
8     context = Context()
9     positions_links = context.get_all_occupation_positions_links()
10    for item in positions_links:
11        if item != None:
12            url = base_url + item
13            urls.append(url)
14
15    process = CrawlerProcess()
16    process.crawl(PositionSpider, urls = urls)
17    process.start()
18
19
20 def step2_occupation_detail_spider():
21     urls = []
22     base_url = 'https://www.kariyer.net'
23     context = Context()
24     occupation_links = context.get_all_occupation_links()
25     for item in occupation_links:
26         if item != None:
27             url = base_url + item
28             urls.append(url)
29

```

### Illustration 3. 2. Scrapyng Data Models and Step Codes by VS Console

In addition to this, other supplementary data are gathered from İŞKUR (Turkish Employment Agency /equivalent Bureau of Employment in Turkey), Turkey Higher education statistics (YÖK Atlas) and TÜİK (Statistics Bureau). Since March 2021, most of the posted job positions on Kariyer.net has been scraped and organised in a relational database model.

### 3.3. Data Patterns (Relational Model)

The relational data model in LMI differs from other types of survey design procedures for data collection. In the relational model, tables are connected at the prime key level to reach the more dimensional and deeper data layers. For instance, occupations are connected to open positions in the job market or vice versa. Itzink (2016) classify relational database models into different categories of type of data. Here we adopted the model and applied the approach for extracted data from kariyer.net as below. The methodology of analysis based on the cluster data is descriptive statistics of the relational tables extracted from large sets of collection procedure as follows: in the below table, all relational model is described, and data type are defined accordingly.

**Table 3.1. Type of Data and Properties in Kariyer.net for Occupations**

<b>Data type</b>	<b>Var</b>	<b>Relational model</b>	<b>Collection/data type</b>
Position (jobs)	title, company name and city	Position id*	Weekly-posted time
Occupation-major and department	Title, text	Id of position-id of occupation and major	Percentage
Occupation-sector	Title, text	Skill - occupation	Weekly-posted time
Occupation skill	Title, text	Occupation id-skill id	Weekly-posted time
Occupation-University	Title, text	Occupation id-university id	Weekly-posted time
Occupation (group of jobs) details	Positions title posted in 2021	Position id	Lowest and highest salary (survey), gender (per cent), participants (gender)

The first table shows the data type and variables as set gathered by Scraping tools.

**Table 3.2. Type of Data for Relational Modeling**

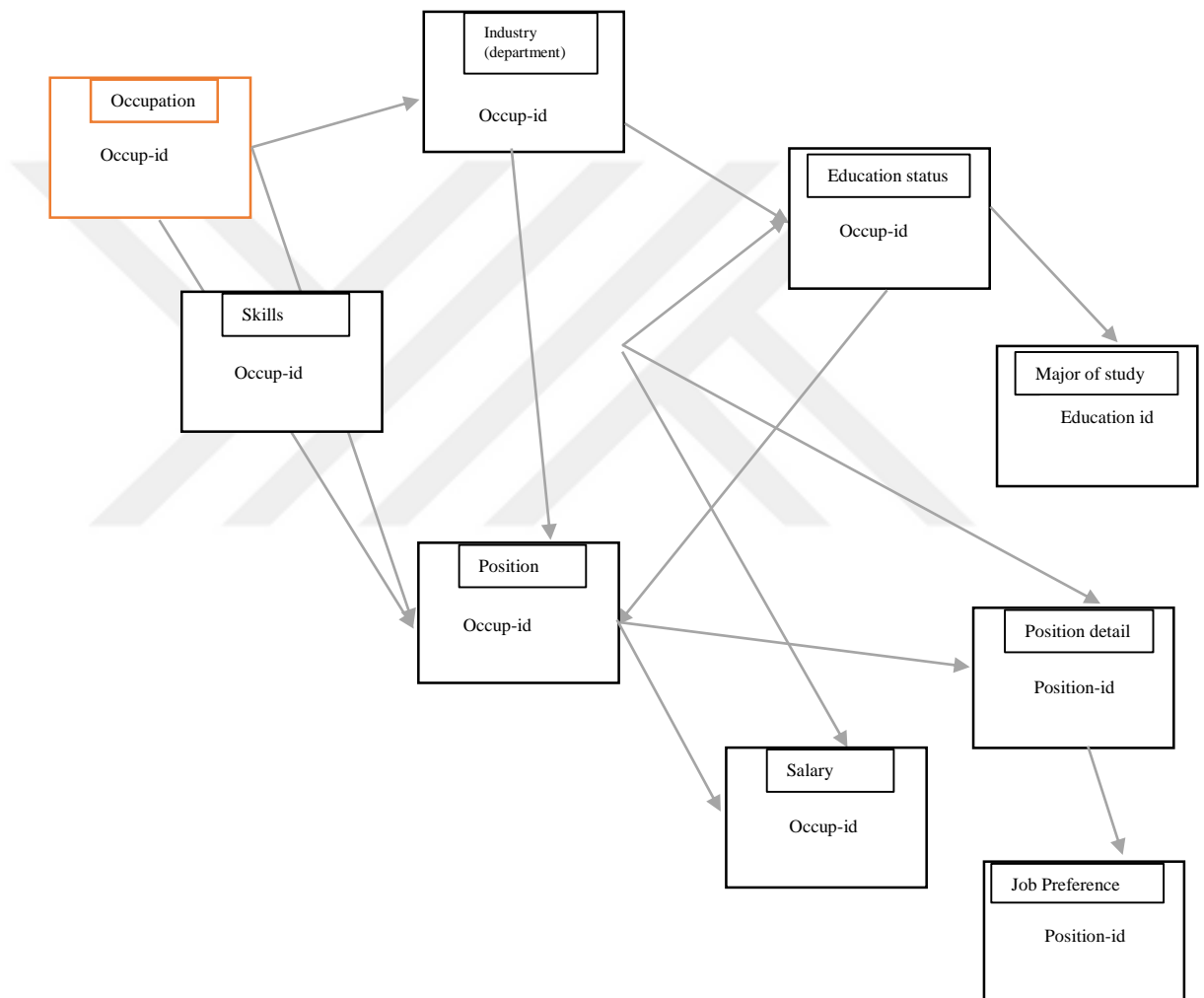
<b>Data type</b>	<b>var</b>	<b>(Prime key)</b>	<b>Date modified</b>
Education	Majors, fields of study and training	Major id, University id	2017-2018 (Undergrads) 2020-2021 (Graduate) Vocational training
People data	Total of undergrad and grad students	Status id (grad or failed)	2021-2022
Cities	Universities cities	Province id	2021-2022
Major	Major names	Occupation id	2021-2022
Skills	Syllabus titles	Skill id *	2021-2022

### 3.4. Data Analysis

The final online transactional processing (OLTP) is prepared in PostgreSQL to create a data warehouse and extract insights from labour market information. Although this table does not reflect all possible data in the market, it is a way to consider how separated sources should be relational to take advantage of it. SQL-based design databases mainly use the relational model.

Given the schematic view of the relational model of tables proposed in Fig 3.1. it should be highlighted how different variables can be connected in a large set by prime keys or any interrelated variable to take meaningful insights into labour market

information (LMI). This way of data collection and applying statistical inferences can decrease the errors of distribution layers of data. It means that more data latency and continuity are attainable in time intervals. The schematic relational model (figure 3.1.) is drawn on the next page. The window function is the technique of joining tables and extracting insight (Itzik Gan, 2016; Itzik Gan, 2019) for calculating rows, moving averages and handling intervals or joint data needs for plotting. Some sample and basic queries are provided in the appendix part as SQL queries.



**Figure 3.1. Relational Model for Data Analysis of Labour Market Information**

The relational model shows how all tables can be connected via occupation id through other resources like skills, salaries, education status etc. After implementing this model, other data like the graduation of educated individuals (supply of labour) and İşkur training were added to compare all insights in one place.

### **3.5. Technical Tools**

#### **3.5.1. Web Scraping by Python Scrape Library**

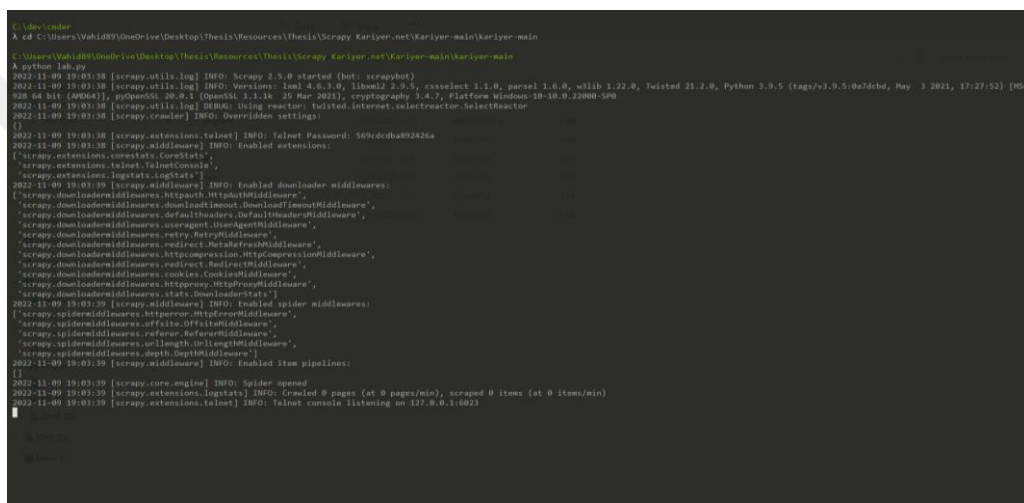
When we started to think about the data collection methodology, the definition of statistical golems (a clay robot from Jewish folklore) caught our interest. Mc Elreath (2020) gives this analogy to define statistical models and discover the truth behind the numbers. There is a need to enrich data to choose an appropriate statistical model. The nature of data can define the parametric assumptions and, after that, statistical inference. One of the opportunities of big data in the labour market is the accessibility and production of data. However, the scale and time dimension are not easy to track, so the process should be automated in a scaled way. To collect data during the thesis, a web scraping methodology is applied. Web scraping is the process in which the access data can be gathered by sending requests to HTML of web browsers, like reading from cells of the spreadsheet but in an unstructured and automatic way. This process saves time and effort for similar logic of multiple pages. It means that web scraping or web harvesting involves programming techniques to process and gather data through narrowed windows and structure databases, even for a million pages simultaneously (Mitchell, R. (2018)). The process of collecting data in our thesis includes the following:

- i. Retrieving HTML data from the domain of Kariyer.net domains
- ii. Parsing the target information (analysing CSS)
- iii. Storing information in the PostgreSQL server
- iv. View and import data from the gathered database
- v. Descriptive analysis and visualisation.

In addition, Visual Studio console and Cmdr software are used to compose, build, run and keep codes and data.

### 3.5.2. Relational Data Modelling by PostgreSQL

PostgreSQL by the University of California at Berkeley (first release 1986) is an open-source SQL language which facilitates the collection of large sets of data. The object-relational database system helps any researcher to create meaningful logic among different variables. As outlined in the relational model in Chapter 2, in this thesis, we do collection by connecting SQL to Scraping with an automated flow of relational data. PgAdmin is an open-source SQL for the administration and development of data gathering as a database. In the collection data period, version 5.2 is deployed.

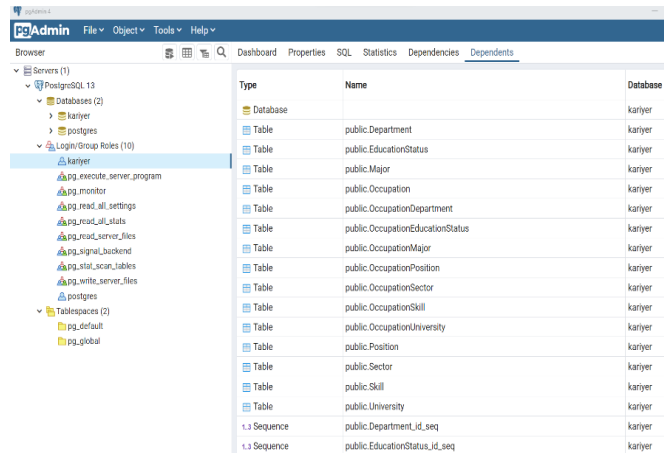


```
C:\Users\Wahid> cd C:\Users\Wahid\OneDrive\Desktop\Thesis\Resources\Thesis\Scrapy_Kariyer.net\Kariyer-main\kariyer-main
Kariyer-main
K python spider.py
2022-11-09 19:03:38 [scrapy.utils.log] INFO: Scrapy 2.5.0 started (bot: scrapybot)
2022-11-09 19:03:38 [scrapy.utils.log] INFO: Versions: lxml 4.6.3.0, libxml2 2.9.5, cssselect 1.1.0, parsel 1.6.0, wlib 1.22.0, Twisted 21.2.0, Python 3.9.5 (tags/v3.9.5:0a7dcd, May 3 2021, 17:27:52) [64-bit AMD64], pyOpenSSL 20.0.1 (OpenSSL 1.1.1k 29 Mar 2021), cryptography 1.4.2, Platform Windows-10-10.0.19044.1506
2022-11-09 19:03:38 [scrapy.utils.log] DEBUG: Using reactor: twisted.internet.selectreactor.SelectReactor
2022-11-09 19:03:38 [scrapy.crawler] INFO: Overridden settings:
{}
2022-11-09 19:03:38 [scrapy.extensions.telnet] INFO: Telnet Password: 569cdc8a092426a
2022-11-09 19:03:38 [scrapy.middleware] INFO: Enabled extensions:
['scrapy.extensions.logstats.LogStats',
 'scrapy.extensions.telnet.TelnetConsole',
 'scrapy.extensions.logstats.LogStats']
2022-11-09 19:03:39 [scrapy.middleware] INFO: Enabled downloader middlewares:
['scrapy.downloadermiddlewares.httpproxy.HttpProxyMiddleware',
 'scrapy.downloadermiddlewares.downloadtimeout.DownloadTimeoutMiddleware',
 'scrapy.downloadermiddlewares.defaultheaders.DefaultHeadersMiddleware',
 'scrapy.downloadermiddlewares.useragent.UserAgentMiddleware',
 'scrapy.downloadermiddlewares.retry.RetryMiddleware',
 'scrapy.downloadermiddlewares.redirect.MetaRefreshMiddleware',
 'scrapy.downloadermiddlewares.httpcompression.HttpCompressionMiddleware',
 'scrapy.downloadermiddlewares.redirect.RedirectMiddleware',
 'scrapy.downloadermiddlewares.cookies.CookiesMiddleware',
 'scrapy.downloadermiddlewares.httpproxy.HttpProxyMiddleware',
 'scrapy.downloadermiddlewares.stats.DownloaderStats']
2022-11-09 19:03:39 [scrapy.middleware] INFO: Enabled spider middlewares:
['scrapy.spidermiddlewares.httperror.HttpErrorMiddleware',
 'scrapy.spidermiddlewares.offsite.OffsiteMiddleware',
 'scrapy.spidermiddlewares.referrer.ReferrerMiddleware',
 'scrapy.spidermiddlewares.urllength.UrlLengthMiddleware',
 'scrapy.spidermiddlewares.depth.DepthMiddleware']
2022-11-09 19:03:39 [scrapy.middleware] INFO: Enabled item pipelines:
[]
2022-11-09 19:03:39 [scrapy.core.engine] INFO: Spider opened
2022-11-09 19:03:39 [scrapy.extensions.logstats] INFO: Crawled 0 pages (at 0 pages/min), scraped 0 items (at 0 items/min)
2022-11-09 19:03:39 [scrapy.extensions.telnet] INFO: Telnet console listening on 127.0.0.1:6023
```

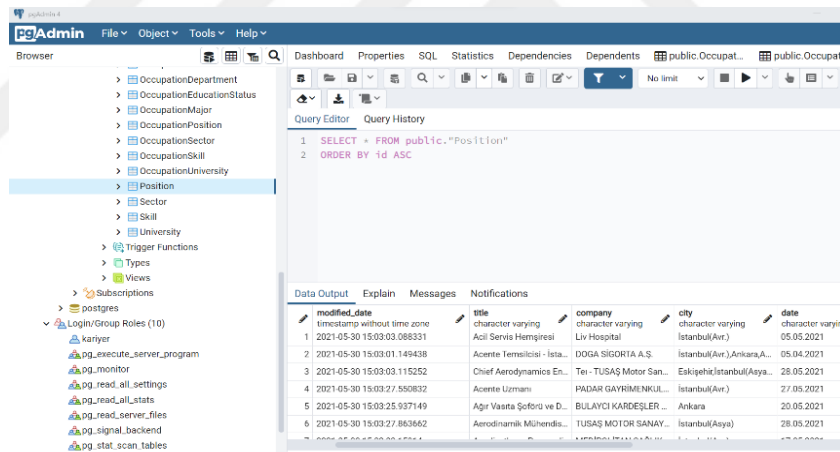
Illustration 3.3. Connection to PostGre to Collect Data

Firstly, we defined a server labelled Kariyer.net. In this process, we defined 15 different tables where the main connecting table (with prime keys) is occupation. With the given id(s), in the occupation table, modified dates, position id, and salaries are defined as outlined in the relational data schematic illustration 3.3. Different tables are assigned to make a bridge and accessibility between different fields of data. For instance, the Occupation table is connected to the department with occupation id and department title, same as education status, major and positions. In the position table, the details of geographic location, company profiles and job titles are defined to import the details of any open job adverts by kariyer.net. As the relational model discussed in

Chapter 2, we skip the explanation and highlight the technology part. The following Figure 3.4 illustrates the schematic of tables, servers, views and sequences.



**Illustration 3. 4. PgAdmin Dashboard for Kariyer.net**



**Illustration 3. 5. Relational Data of Scraped Tables**

### 3.5.3. Extracting Insights and Visualisation

We used Python packages, basic plotting tools in Excel and google sheets, and some geographical visualisation packages to plot data. We used the Rstudio tidyverse package to clean data to check the redundancy or dirty data issues. After finalising the data, we run queries and plot the main insight as provided in chapter 4 of the thesis.

## CHAPTER IV

### RESULTS AND DISCUSSION

#### 4.1. General Insight

From 314980 open job vacancies (the data can differ due to different types of accessibility with Kariyer.net) as our sample size to evaluate the big picture and some initial insight into the Turkey job market from March 2021 to June 2022. The table below shows only the unique job vacancy in terms of industry and job title to rank which industries are leading and how are the career path and its dominance: Using the Python dictionaries to batch the data, we collected the related or semi-related positions into 20 categories which differ that Kariyer.net industry labelling. The result shows that the service industry is leading, and interestingly, manufacturing has a bigger size than IT and Telecom (the main perception is that digital skills are highly demanded after the pandemic). On the other hand, administrative and purely office-like jobs are a threat by diminishing if we consider this sample size.

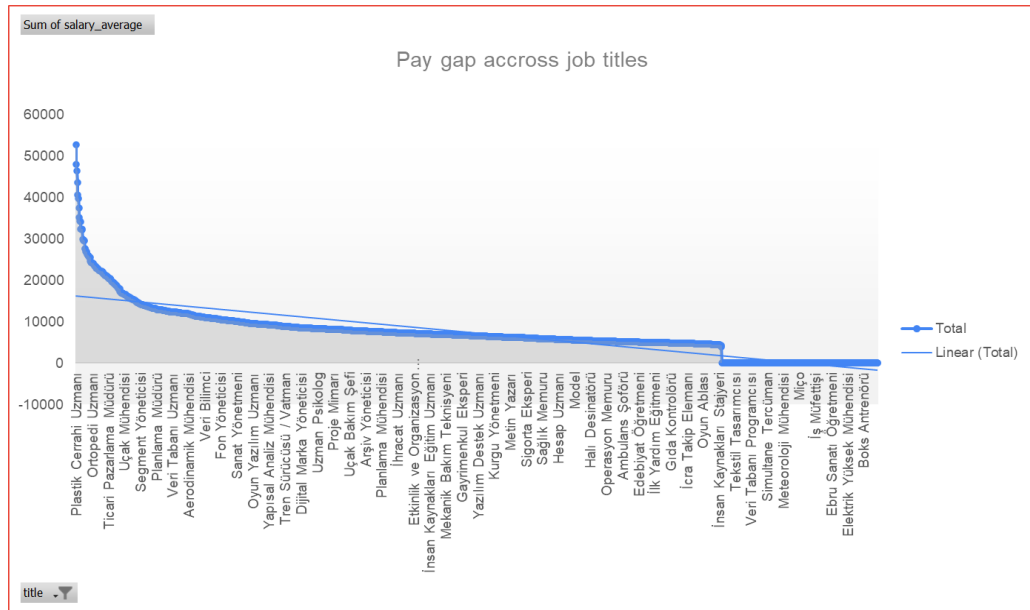


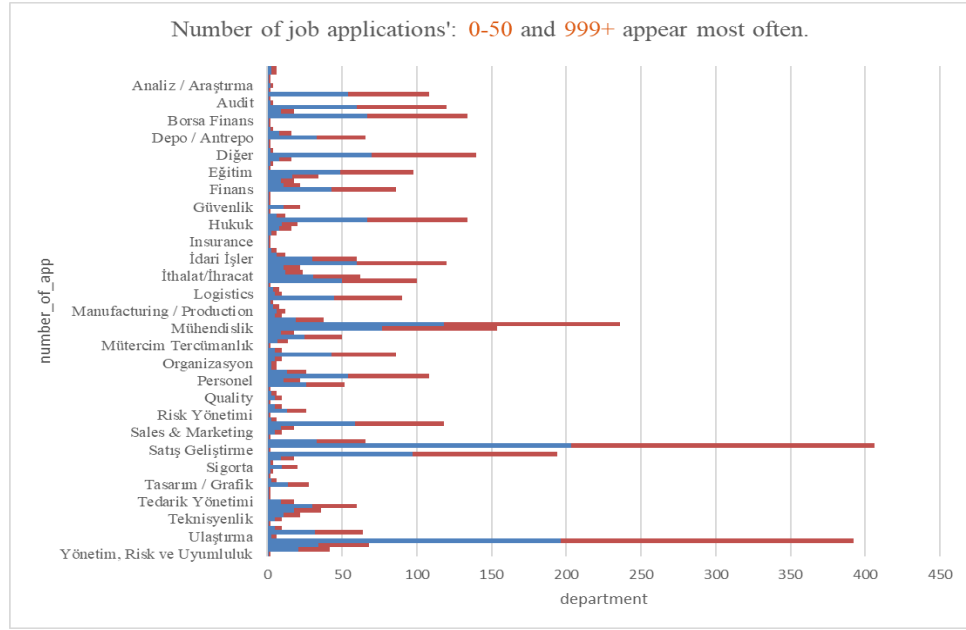
Figure 4.1. Relational Data of Scraped Tables

About all available data, salaries based on occupation and industry show a huge gap in some industries and occupations. The health industry is leading, and surgeon is on higher outlier, but manufacturing, specifically the construction sector, is considered a luxurious job in Turkey. The below graph shows how the salaries vary.

Interestingly the trend of the pay gap for other occupations is flexible. The speciality in sales, marketing and even aviation is sometimes near. The mean of salaries is calculated through 1600 available data out of 2670. It means that more than 1000 occupations are not listed in this category due to limitation of participation or nullity of data. To have a robust analysis, the participation of over +50 people was selected to reflect the reality of the job market. For example, in Textile, Technician jobs and some technical jobs like data analysis, the participation is low, or the nullity of data is most common.

#### **4.2. Preference of Job Market (Supply Side)**

The graph shows how the market can create employment across different sectors and departments. The data is aggregated analysis with preference ranging from 0-50 up to over 999, meaning that to what extent different people have applied for the positions within the sectors. Although there would be multiple applications from a single person, we can conclude that sales growth positions created more job opportunities and more applications. According to Figure 4.2. blue bars means the aggregate index of created jobs per position and, on the same scale, the applications or preference of people to be employed. Management is the most desired sector, while analysis job has the median shortage and surpasses. Turkish labour market calls for more vendors and manufacturing during 2021, relying only upon kariyer.net data.



**Figure 4.2. Employment Demanding (Open Positions to Apply)**

**Table 4.1. Aggregate Data for Employment Demand**

No.	Sector/Industry	Turkish Labels	Unique Open Job Vacancy	Rank	Total
1	Service	Hizmet Danışmanlık Güvenlik	9306 3324 529	1	13159
2	IT and Telecom	Bilişim Telekomunikasyon	5264 2434	3	7696
3	Construction	Yapı	5228	5	5228
4	Health	Sağlık	4708	7	4708
5	Fashion	Tekstil Mobilya ve aksesuar	4359 374	6	4733
6	Agriculture and Food	Gıda Tarım / Ziraat Hayvancılık	3989 247 130	8	4366
7	Manufacturing*	Üretim / Endüstriyel Otomotiv Ev Eşyaları Hızlı tüketim malları Dayanaklı Tüketim Ürünleri	3825 3368 620 252 14	2	8079
8	Education	Eğitim	3488	10	3488
9	Finance, Banking and Insurance	Finans – Ekonomi Sigortacılık	3179 2166	4	5344
10	Entertainment	Turizm	2678	12	2678

**Table 4.1. (cont.)**

11	Electronics	Elektrik & Elektronik	2667	13	2667
12	Mining	Maden ve Metal Sanayi	2562	14	2562
13	Energy	Enerji	2299	16	2299
14	Transportation	Taşımacılık Havacılık Denizcilik	1995 590 556	11	3141
15	Chemistry	Kimya	1863	17	1863
16	Commerce and retail	Ticaret Perakende	1704 828	15	2532
17	Media	Reklam ve Tanıtım Medya Basım yayın İletişim danışmanlığı	1672 1100 780 81	9	3633
18	Entertainment	Eğlence - Kültür – Sanat	1208	18	1208
19	Nature	Çevre Atık Yönetimi ve Geri Dönüşüm Orman Ürünleri Kaynak ve kesme ekipmanları	599 96 86 41	20	795
20	Administrative	Topluluklar Arşiv yönetimi ve Saklama Ofis / Büro Malzemeleri	595 187 54	19	836
		<b>Grand Total</b>			<b>81045</b>

### 4.3. Employment Status (Demand Side)

Table 4.1. demonstrates the sectors where unique positions have been opened in the past year. From 314 thousand, the unique job titles are 81 thousand, which shows the market is experiencing a shift to create new jobs. It means that 25% of the labour market consists of new demands requiring new skill sets, experience and competencies. In the education part, we will highlight perquisites, but service, manufacturing, and IT are the most dominant employment hubs and attractive job adverts, placing 1 to 3 in the ranking. The data trend shows that Turkey requires demanding salespeople in different fields or occupations. The next most frequent titles are related to production. Data interestingly says how one should reconsider about bachelor syllabus. Graph 4.4. shows that 23 jobs are not specifically mentioned the required degree. Interestingly high school is more demanding than a bachelor's in this context. If people have higher education, though, a master's degree is a better option to meet the requirement.

## FREQUENCY (JOB TITLES)

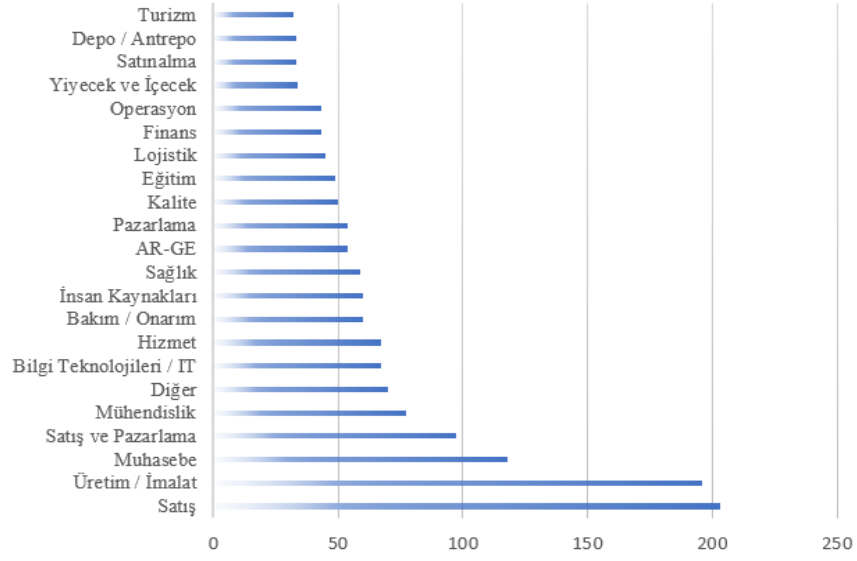


Figure 4.3. Job Titles Frequency on Unique Titles in Different Sectors

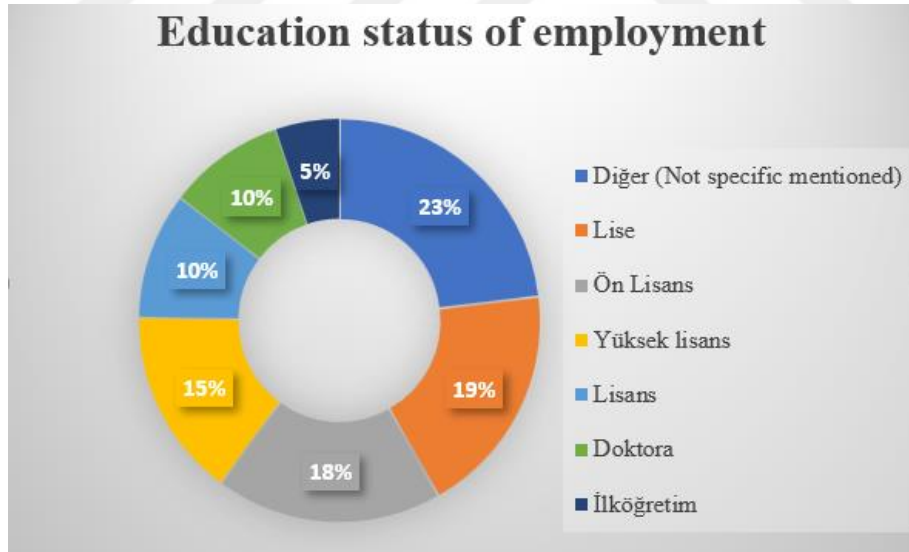


Figure 4.4. Education and Training Needs of Labour Market

**Table 4.2. Number of Undergraduate Students Registered for 2017-18, Turkey**

2017-2018	Dep. No.	New admitted			Registered		
		M	W	T	M	W	T
Faculties/departments	1,497	464,379	495,259	959,638	294,2089	276,8481	5,710,570
4 Year higher vocational education	339	18,247	17,050	35,297	91,373	82,998	174,371
2-year vocational education (associate)	913	122,085	117,568	239,653	683,446	442,211	1,125,657
<b>Total</b>	<b>2,749</b>	<b>604,711</b>	<b>629,877</b>	<b>1,234,588</b>	<b>3,716,908</b>	<b>3,293,690</b>	<b>7,010,598</b>

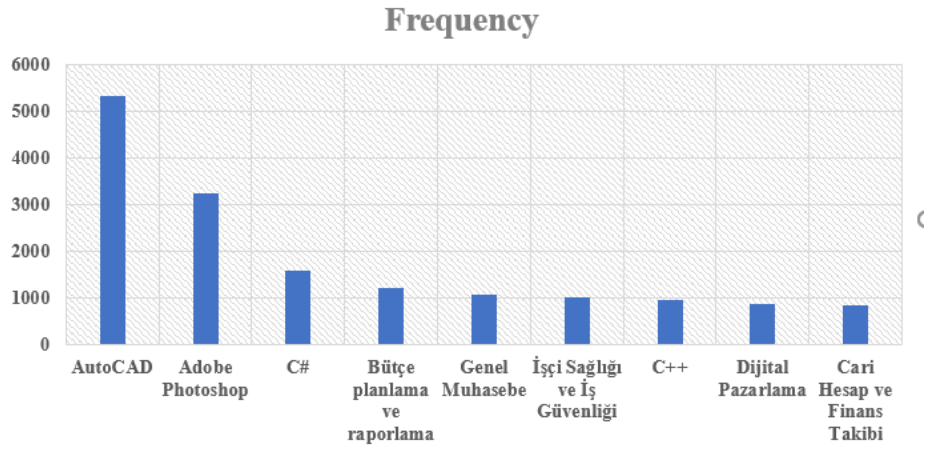
**Table 4.3. Number of Graduate Students Registered for 2020 -21, Turkey**

2020-2021	Dep. No.	New admitted			Registered		
		M	W	T	M	W	T
Master		81,918	69,781	151,699	177,977	165,592	343,569
PhD		9,552	9,324	18,876	55,558	50,590	106,148
<b>Total</b>	<b>569</b>	91,470	79,105	170,575	233,535	216,182	<b>449,717</b>

**Table 4.4. Number of Vocational Training by Iskur, Jan-Dec 2021**

Educational status	Vocational Training			Training of İşbaşı		
	Men	Women	Total	Men	Women	Total
<b>Total</b>	<b>28,573</b>	<b>72,928</b>	<b>101,501</b>	<b>176,313</b>	<b>183,857</b>	<b>360,170</b>
Illiterate	43	74	117	253	107	360
Literacy (writing and reading)	1,530	6,330	7,860	5,379	5,259	10,638
Primary school	11,421	35,145	46,566	64,594	55,215	119,809
Highschool	10,606	19,291	29,897	68,829	62,291	131,120
Associate degree	2,575	6,822	9,397	18,223	31,187	49,410
Bachelor	2,349	5,178	7,527	18,592	29,134	47,726
Master	46	86	132	426	652	1,078
PhD	3	2	5	17	12	29

The main point is that many jobs need to target specifically education status. However, higher education is most welcome, according to the data of Yök (2018, 2021) given in Tables 4.2. and 4.3. with over 7.5 million people. This size of educated people imposes the expectation to fill the most deserving seats as outlined in a queue of buses in the literature review, which leads to an imbalance in shortage and surpasses both sides of the labour market. Employers who want to hire may face high expectations, and in return, people may stay unemployed as they expect better jobs. Still, quality matters and the industry expect high skills in sales, accountants and finance-savvy alums. Digital marketing Security and health care are in the ranking as well.



**Figure 4.5. Skill Frequency Across Positions**

The last table showed the ranking of universities when participants responded when they got their job from the platform. Interestingly universities commonly not listed as ivy league of Turkey are among the top employed alums. There is no apple-to-apple comparison as Istanbul University is one of the mother universities in Turkey. For example, Anadolu University is top among others as first rank; however, Kocaeli University is top for 4 and 5 most recruited alums in all positions. We must anchor and rely on something other than this data as aggregated and comprehensive. However, it still values considering the overall performance of departments and open fields in the universities to reconsider the training and higher education for the labour market. Finally, in Chapter 5, we will make suggestions and conclude the findings.

## CHAPTER V

### CONCLUSIONS

#### 5.1. Conclusion

It is recommended that a local model boost the data and statistical analytics like a semi-Bureau of Labour statistics if it is not a must but can be strongly recommended. Turkey experiences a very harsh inflation rate, and since 2005, this year, there have been two round adjustments in minimum salaries, affecting all employment parameters. Overall population variables might also impact the unemployment rate, access to education and welfare indices. Hence, the key mechanism to overcome the subtasks and make the privilege of inherent opportunities is creating a special office to track live data of Turkey to project the dynamics and plan accordingly. On the other hand, this situation encourages people to immigrate or work remotely, which can disrupt and create a gap in talent shortage. To have all this on one page, the conventional methodology of Tük, işkur etc., to grasp data impose limitations.

Our study shows that the common hypothesis is quite different from that generalisation of employment rate in dominant industries. The main problem is the syllabus of the university, where the industry expects a strong background in marketing, sales and manufacturing skilled graduates. Turkey, at least in recent years, welcomed more sales specialists than other occupations if it is confined to this data. Vocational training is very underrated in Turkey. It holds only 25 per cent of education for graduates and faculties. Data from İşkur shows that people can be trained for some sectors/occupations with real cases, saving financial and time for job seekers.

#### 5.2. Practical Recommendations

Data is good but precious when it leads to the right insights. As here we speak with descriptive analysis, generalisation is very limited, but there are some important

takeaways. Firstly, the data pattern demonstrates that there is no universal occupation classification in Turkey's employment ecosystem, or if there is, non-systematic. This classification can boost the nature of the labour market in Turkey to plan for education, salary adjustment and standardisation of jobs.

Second, we can consider the location, job dynamics and skill factor for salary adjustment. This can help balance and motivate people, which ends in productivity. Anchoring this type of data helps leverage minimum wage from multiple perspectives. For instance, diagnosing vulnerable industries, sectors, and employers would be possible. On the other hand, the budget, planning, and even opening the education and training can be regulated and adjusted with the insightful mind to save time and financial costs. In addition, the government's support package for social security, tax, or any endowment for creating jobs and promoting employment will be trackable. The questions like did they employ individuals or whether the supporting package or credits are utilised in a promised way or not are easy to find. Geographical data can help the government to adjust internal or external immigration across the country. Government can guide individuals with live data to show which locations can create more job opportunities, or if one place tolerates surplus or shortage, then actions can be implemented accordingly.

Lastly, the government or any private or state stakeholder should invest in the real-time labour data to adjust and take actions for empowering individuals or alums at the right time, check the needs of the market and education and find the mismatch and lastly, insightful planning for short-term crises or uncertainties.

### **5.3. Suggestions for Future Studies**

Based on the research scope and time and resource scarcity, some topics as listed below can be carried out for the following research or case studies:

- Collect more data from multiple resources like LinkedIn, Glassdoor and İşkur and compare the data to check the validity.

- Working on clustering of educational data to check the mismatch and vulnerability of education
- Geographical or spatial analysis to deep dive into the mobility of the market.
- Cross-check education or syllabus to check whether they align with employment as competency-based education and planning.

#### **5.4. Limitations of Thesis**

Although the 314 thousand job positions, any figures or dataset from the main source is inevitable as this thesis is limited to openly accessible data; therefore, it is possible to encounter any difference among the tables and figures regarding salaries, education status, job positions, etc. Although the research endeavours to tide data in a very disciplined way, the errors of collection and, therefore, the descriptive analysis is not zero. The other primary limitation is that this collected data only reflects the 2021-22 years and only some of the year's open positions; subsequently, main employment and job adverts are announced on different platforms or mediums, which the thesis cannot reflect all markets. Moreover, illegal or even ignoring paper-based job announcements can affect the overall real data of the labour market.

- Data from thesis collection can differ from published data of kariyer.net (as kariyer.net updates all-time salaries etc.).
- Lack of time to extract and define statistical inference to boost the meaningfulness of data.
- Lack of time and resources to access other mediums or sources of data like işkur, municipalities, ministry of labour or social security organisation.

## REFERENCES

- Akçomak, S., & Gürcihan, H. B. (2013). *Türkiye işgücü piyasasında mesleklerin önemi: hizmetler sektörü istihdamı, işgücü ve ücret kutuplaşması*. Türkiye Cumhuriyeti Merkez Bankası Çalışma Tebliği, 13, 21.
- Akçoraoğlu, A. (2010). *Employment, economic growth, and labour market performance: The case of Turkey*. *Ekonomik Yaklaşım*, 21(77), 101-114.
- Aksu, E., Erzan, R., & Kırdar, M. G. (2022). *The impact of mass migration of Syrians on the Turkish labour market*. *Labour Economics*, 102183.
- Aslan, G. (2019). *Türkiye’de asgari ücretli çalışan sayısı ve ücret seviyelerinin değişimi (2003-2017 hanehalkı işgücü anketleri veri analizi*. *SGD-Sosyal Güvenlik Dergisi*, 9(1), 141-159.
- Ben-Gan, I. (2016). *T-Sql Fundamentals*. Microsoft Press, Third edition, ISBN 978-1-5093-0200-0
- Ben-Gan, I. (2019). *T-SQL Window Functions*. Microsoft Press.
- Bhattacharya, T., Bhandari, B., & Bairagya, I. (2020). *Where are the jobs? Estimating skill-based employment linkages across sectors for the Indian economy: an input-output analysis*. *Structural Change and Economic Dynamics*, 53, 292-308.
- Biffignandi, S., & Bethlehem, J. (2012). *Web surveys: methodological problems and research perspectives*. *Advanced statistical methods for analysing large data sets*, pp. 363–373.
- Del Carpio, X. V., & Wagner, M. C. (2015). *The impact of Syrian refugees on the Turkish labour market*. World Bank policy research working paper, (7402).
- Einav, L., & Levin, J. (2014). *Economics in the age of big data*. *Science*, 346(6210), 1243089.
- European Commission. M. Mirigliano (2022) *Digital Jobs and Skills Platform: 2022 State of the Union: "2023 will be the European year of Skills*. Short link:<https://bit.ly/3MhKSuO>
- Garaso, S., Steelman, C., Djumaliye Y., (2021). *NESTA Project: Boosting labour market intelligence for local decision-makers: A new method to measure regional skill demand across the UK*: <https://www.nesta.org.uk/project-updates/boosting-labour-market-intelligence-local-decision-makers/>
- Gündüz-Hoşgör, A., & Smits, J. (2008, March). *Variation in labour market participation of married women in Turkey*. In *Women's Studies International Forum* (Vol. 31, No. 2, pp. 104-117). Pergamon.

- Harrington, J. L. (2016). *Relational database design and implementation*. Morgan Kaufmann.
- Ilkcaracan, I., & Selim, R. (2007). *The gender wage gap in the Turkish labour market*. *Labour*, 21(3), 563-593.
- International Labor Organization (2020), *The feasibility of using big data in anticipating and matching skills needs* – International Labour Office – Geneva. ISBN 978-92-2-032854-5 (web PDF)
- Işık, V. (2016). *Türkiye’de genç işsizliği ve genç nüfusta atalet*. *Hak İş Uluslararası Emek ve Toplum Dergisi*, 5(11), 130-145.
- Johnson, E. (2016). *Can big data save labour market information systems?* RTI Press. RTI Press Policy Brief No. PB-0010-1608 <https://doi.org/10.3768/rtipress.2016.pb.0010.1608>
- Kariyer.net. Politika lab (2022). *İstihdam Etki Raporu*, <https://tr.kariyer.net/etki-raporu/>
- Khaouja, I., Mezzour, G., Carley, K. M., & Kassou, I. (2019). *Building a soft skill taxonomy from job openings*. *Social Network Analysis and Mining*, 9(1), 1-19.
- Leontief, W. (Ed.). (1986). *Input-output economics*. Oxford University Press.
- LinkedIn Global Trend Report (2022). *Data-driven insights into the changing world of work*, October 2022: <https://business.linkedin.com/talent-solutions/global-talent-trends>
- Loayza, N. V., Ulysea, G., & Utsumi, T. (2018). *Informality and the labour market effects of mass migration: Theory and evidence from Syrian refugees in Turkey*. *In the 7th Turkey labour market network meeting*. Accessed September (Vol. 9, p. 2020).
- Maass, W., Parsons, J., Purao, S., Storey, V. C., & Woo, C. (2018). *Data-driven meets theory-driven research in the era of big data: Opportunities and challenges for information systems research*. *Journal of the Association for Information Systems*, 19(12), 1.
- Mahmut, Ö. R., & Suna, H. E. (2020). *The linkage between vocational education and the labour market in Turkey: Employability and skill mismatch*. *Kastamonu Eğitim Dergisi*, 28(2), 558-569.
- McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.
- Mitchell, R. (2018). *Web scraping with Python: Collecting more data from the modern web*. O'Reilly Media, Inc.

- Mrsic, L., Jerkovic, H., & Balkovic, M. (2020, March). *Interactive skill-based labour market mechanics and dynamics analysis system using machine learning and big data*. In Asian Conference on Intelligent Information and Database Systems (pp. 505-516). Springer, Singapore.
- Mülteciler Derneği. (2022). *Türkiye'deki Suriyeli Sayısı*: <https://bit.ly/3TrFB5H>
- Sevinc, D., Green, A., Bryson, J. R., Collinson, S., Riley, R., & Adderley, S. (2020). *Ensuring skills are available in the right locations: Are we there yet? A regional analysis of qualification gaps*. *Regional Studies*, 54(8), 1149–1159.
- Tansel, A. and Acar, E.Ö. (2017). *Labor mobility across the formal/informal divide in Turkey: Evidence from individual-level data*. *Journal of Economic Studies*, Vol. 44 No. 4, pp. 617–635. <https://doi.org/10.1108/JES-06-2015-0103>
- Tutar, K. (2015). *Türkiye işgücü piyasasında iş arama kanallarının niteliği*. *İş ve Hayat*, 1(2), 93-106.
- Türkiye İstatistik Kurumu (2022). *Labour Force Statistics, September 2022*. <https://data.tuik.gov.tr/Bulten/Index?p=Labour-Force-Statistics-September-2022-45655>
- PostgreSQL, B. (2022). *PostgreSQL*. Web resource: [http://www. PostgreSQL.Org/about](http://www.PostgreSQL.Org/about).
- World Economic Forum (2022). *Future of work agenda*. <https://www.weforum.org>
- Yunculer, H. B. G., & Yunculer, C. (2016). *Minimum Wage Effects on Labor Market Outcomes in Turkey* (No. 1614).

# APPENDIXES

## APPENDIX A

In this Appendix, the written codes and syntax to collect data are given below, as mentioned in the logic of data collection in Chapter 3. For writing codes, python spider and scraping tools were used to connect and create a database on PostgreSQL. The procedure started with Scraping, then a relational database running the queries to catch insights. In addition to codes, versions, library detail and queries are listed.

### Python 3.1.1

#### Spider Library

Code IDE: VS (Microsoft)

#### A) Lab.py Codes:

```
From sqlalchemy, import Column, String, BigInteger
from sqlalchemy.sql.sqltypes import DateTime, Integer, Float
from sqlalchemy.ext.declarative import declarative_base
Base = declarative_base()
class Occupation(Base):
    __tablename__ = 'Occupation'
    id = Column(BigInteger(), primary_key=True, autoincrement=True)
    modified_date = Column(DateTime())

    posid = Column(String())
    title = Column(String())

    salary_average = Column(Float())
    salary_lowest = Column(Float())
    salary_highest = Column(Float())
    salary_people = Column(Integer())

    gender_male = Column(Float())
```

```

gender_female = Column(Float())
gender_people = Column(Integer())

link = Column(String())
positions_link = Column(String())

def __repr__(self):
    return "Occupation: (title='%s', id='%s', link='%s', salary_people='%s')" % (
        self.Title, self.id, self.link, self.salary_people)

```

```

class Skill(Base):
    __tablename__ = 'Skill'

    id = Column(BigInteger(), primary_key=True, autoincrement=True)
    modified_date = Column(DateTime())

    title = Column(String())

```

```

class OccupationSkill(Base):
    __tablename__ = 'OccupationSkill'

    id = Column(BigInteger(), primary_key=True, autoincrement=True)
    modified_date = Column(DateTime())

    occupation_id = Column(BigInteger())
    skill_id = Column(BigInteger())

    title = Column(String())

```

```

class University(Base):
    __tablename__ = 'University'

    id = Column(BigInteger(), primary_key=True, autoincrement=True)
    modified_date = Column(DateTime())

```

```
title = Column(String())
```

```
class OccupationUniversity(Base):
```

```
    __tablename__ = 'OccupationUniversity'
```

```
    id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
    modified_date = Column(DateTime())
```

```
    occupation_id = Column(BigInteger())
```

```
    university_id = Column(BigInteger())
```

```
title = Column(String())
```

```
order = Column(Integer())
```

```
class EducationStatus(Base):
```

```
    __tablename__ = 'EducationStatus'
```

```
    id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
    modified_date = Column(DateTime())
```

```
title = Column(String())
```

```
class OccupationEducationStatus(Base):
```

```
    __tablename__ = 'OccupationEducationStatus'
```

```
    id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
    modified_date = Column(DateTime())
```

```
    occupation_id = Column(BigInteger())
```

```
    educationstatus_id = Column(BigInteger())
```

```
title = Column(String())
```

```
order = Column(Integer())
```

```
class Major(Base):
```

```
__tablename__ = 'Major'
```

```
id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
modified_date = Column(DateTime())
```

```
title = Column(String())
```

```
class OccupationMajor(Base):
```

```
__tablename__ = 'OccupationMajor'
```

```
id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
modified_date = Column(DateTime())
```

```
occupation_id = Column(BigInteger())
```

```
major_id = Column(BigInteger())
```

```
title = Column(String())
```

```
percentage = Column(Float())
```

```
class Sector(Base):
```

```
__tablename__ = 'Sector'
```

```
id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
modified_date = Column(DateTime())
```

```
title = Column(String())
```

```
class OccupationSector(Base):
```

```
__tablename__ = 'OccupationSector'
```

```
id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
modified_date = Column(DateTime())
```

```
occupation_id = Column(BigInteger())
```

```
sector_id = Column(BigInteger())
```

```
title = Column(String())
```

```
percentage = Column(Float())
```

```
class Department(Base):
```

```
    __tablename__ = 'Department'
```

```
    id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
    modified_date = Column(DateTime())
```

```
    title = Column(String())
```

```
class OccupationDepartment(Base):
```

```
    __tablename__ = 'OccupationDepartment'
```

```
    id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
    modified_date = Column(DateTime())
```

```
    occupation_id = Column(BigInteger())
```

```
    department_id = Column(BigInteger())
```

```
    title = Column(String())
```

```
    percentage = Column(Float())
```

```
class Position(Base):
```

```
    __tablename__ = 'Position'
```

```
    id = Column(BigInteger(), primary_key=True, autoincrement=True)
```

```
    modified_date = Column(DateTime())
```

```
    title = Column(String())
```

```
    company = Column(String())
```

```
    city = Column(String())
```

```
date = Column(String())
link = Column(String())
general = Column(String())
criteria = Column(String())
information = Column(String())
statistics = Column(String())
```

```
class OccupationPosition(Base):
    __tablename__ = 'OccupationPosition'

    id = Column(BigInteger(), primary_key=True, autoincrement=True)
    modified_date = Column(DateTime())

    occupation_id = Column(BigInteger())
    position_id = Column(BigInteger())
```

The connection between kariyer.net and PostgreSQL

```
from models import Base, Occupation
from sqlalchemy import create_engine
from sqlalchemy.orm import sessionmaker
from datetime import datetime
from typing import List, TypeVar

class Context():

    T = TypeVar('T')

    def __init__(self):
        connection_string = 'postgresql://kariyer:kariyer@localhost:5432/kariyer'
        engine = create_engine(connection_string, max_overflow=-1, echo = False)
        Base.metadata.create_all(engine)
        self.Session = sessionmaker(bind=engine)

    def add(self, model: T) -> T:
```

```
now = datetime.utcnow()
model.modified_date = now
```

```
Session = self.Session()
session.add(model)
session.commit()
session.refresh(model)
session.close()
return model
```

```
def add_many(self, models):
    now = datetime.utcnow()
    For the model in models:
        model.modified_date = now

    session = self.Session()
    session.add_all(models)
    session.commit()
    session.close()
```

```
def update(self, model: T, id, dict):
    now = datetime.utcnow()
    dict.update({model.modified_date: now})

    session = self.Session()
    session.query(model).filter_by(id = id)\
        .update(dict, synchronize_session = False)
    session.commit()
    session.close()
```

```
def get_by_id(self, model: T, id) -> T:
    session = self.Session()
    item = session.query(model).filter_by(id = id)
    session.close()
```

```

return item[0]

def get_by_title(self, model: T, title) -> List[T]:
    session = self.Session()
    items = session.query(model).filter_by(title = title)
    session.close()
    return items

def get_by_like_title(self, model: T, title) -> List[T]:
    session = self.Session()
    items = session.query(model)\
        .filter(model.title.like('% {title} %'.format(title = title)))
    session.close()
    return items

def get_by_link(self, model: T, link) -> List[T]:
    session = self.Session()
    items = session.query(model).filter_by(link = link)
    session.close()
    return items

def get_all(self, model: T) -> List[T]:
    session = self.Session()
    items = session.query(model).all()
    session.close()
    return items

def get_all_distinct_titles(self, model: T) -> List[str]:
    session = self.Session()
    items = session.query(model.title).distinct().all()
    session.close()
    return [x[0] for x in items]

def get_all_occupation_links(self) -> List[str]:

```

```
session = self.Session()
items = session.query(Occupation.link).all()
session.close()
return [x[0] for x in items]
```

```
def get_all_occupation_positions_links(self) -> List[str]:
    session = self.Session()
    items = session.query(Occupation.positions_link).all()
    session.close()
    return [x[0] for x in items]
```

```
def get_occupation_id_by_positions_link(self, link) -> Occupation:
    session = self.Session()
    item = session.query(Occupation.id).filter_by(positions_link = link)
    session.close()
    return item[0][0]
```

## **B) Data Insight (PostgreSQL) Sample Queries**

```
SELECT *
From "public" to "Occupation."
WHERE "salary_average" is not NULL
ORDER BY "salary_average" DESC -- ASC
LIMIT 15
```

```
SELECT "public"."Occupation".id as "Occupation Id",
       "public"."Occupation".title as "Occupation Title",
       "public"."OccupationMajor".title as "Major Title"
From "public.Occupation."
INNER JOIN "public"."OccupationMajor" ON
         "public"."Occupation".id = "public"."OccupationMajor".occupation_id;
```

```
-----
SELECT *
```

```
--SELECT COUNT(*)
--SELECT COUNT (DISTINCT "title")
From "public"."OccupationSkill."
--WHERE "title" = 'C#'
--SELECT *
SELECT COUNT(*)
--SELECT COUNT (DISTINCT "title")
From "public"."Position."
--WHERE "date" like '%.02.2021.'
--WHERE "city" = 'Ankara'
--WHERE "city" like '%Ankara%.'
--WHERE "date" like '%.03.2021' and "city" = 'Ankara'
```

## APPENDIX B

We used multiple resources like Excel and Python packages to work on data visualisation. The extra table are provided as below. In the chapter 4, main tables are plotted and these are as complimentary data to support our findings.

**Table B.1. Industry Ranking Based on Collected Data**

<b>Industry</b>	<b>Frequency (Job Titles)</b>
Satış	203
Üretim / İmalat	196
Muhasebe	118
Satış ve Pazarlama	97
Mühendislik	77
Diğer	70
Bilgi Teknolojileri / IT	67
Hizmet	67
Bakım / Onarım	60
İnsan Kaynakları	60
Sağlık	59
AR-GE	54
Pazarlama	54
Kalite	50
Eğitim	49
Lojistik	45
Finans	43
Operasyon	43
Yiyecek ve İçecek	34
Satınalma	33
Depo / Antrepo	33
Turizm	32

**Table B.2. Education Status Rankings**

<b>Rank</b>	<b>Degree</b>	<b>Ranking Percentage</b>
1	Diğer (Not specifically mentioned)	23,12%
2	Lise	18,67%
3	Ön Lisans	18,28%
4	Yüksek lisans	15,18%
5	Lisans	10,15%
6	Doktora	9,44%
7	İlköğretim	5,17%
	<b>Grand Total</b>	<b>100,00%</b>

**Table B.3. Demanding Skills Rankings**

<b>Most required Skills</b>	<b>Frequency</b>
AutoCAD	5311
Adobe Photoshop	3238
C#	1586
Bütçe planlama ve raporlama	1214
Genel Muhasebe	1062
İşçi Sağlığı ve İş Güvenliği	1022
C++	951
Dijital Pazarlama	861
Cari Hesap ve Finans Takibi	834
Finans	834

# CURRICULUM VITAE

## Personal Information:

Name - Surname: Vahid Farajjoheddar

## Education:

2020-2022 MA in Economics, Ibn Haldun University, Turkey

2013-2015 MA in Human Resources, University of Tehran, Iran

## Experience:

2010 – Present ..... Kariyer.net

2010 – 2012 ..... WorqCompany