CHAPTER-1

ABOUT THE COMPANY

Fig.1.1 Company Image

1.1 History

KPMG, short for Klynveld Peat Marwick Goerdeler, is one of the "Big Four" global accounting and professional services firms. It provides a wide range of services including audit, tax, and advisory services to various clients, including corporations, governments, and non-profit organizations.

KPMG's history dates back to 1870 when Piet Klynveld founded an accounting firm in Amsterdam, Netherlands. In 1917, William Barclay Peat and James Marwick established their own accounting firms in London and New York respectively. These firms grew and expanded their services over the years.

In 1987, Klynveld Kraayenhof & Co.(K) merged with Peat Marwick Mitchell (PM) to form KPMG. The merger brought together KPMG's international presence and Peat Marwick's strong U.S. practice. The addition of Ernst & Whinney in 1989 further solidified KPMG's global position.

Throughout the 1990s and 2000s, KPMG continued to expand its services and global reach. It established a strong presence in emerging markets and extended its offerings beyond traditional accounting services to include advisory and consulting services. Like other major accounting firms, KPMG faced challenges in the early 2000s due to accounting scandals such as Enron and WorldCom. These scandals led to increased regulatory oversight and changes in the accounting industry, affecting how firms like KPMG operated.

Over the years, KPMG diversified its services to include not only traditional auditing and tax services but also management consulting, technology consulting, risk management, and more. This diversification aimed to provide clients with comprehensive solutions to complex business challenges.

In line with global trends, KPMG also emphasized its commitment to sustainability and corporate responsibility. The firm began offering services related to environmental, social, and governance (ESG) issues to help organizations navigate the evolving landscape of responsible business practices.

As of my last update in September 2021, KPMG continued to be a major player in the accounting and professional services industry. It adapted to technological advancements and changing client needs, positioning itself as a comprehensive service provider in a rapidly evolving business landscape.

1.2 Company Services

KPMG provides independent auditing and assurance services to help businesses ensure the accuracy and reliability of their financial statements. This helps build trust among stakeholders and meet regulatory requirements. KPMG offers comprehensive tax planning, compliance, and advisory services to help clients manage their tax obligations efficiently and effectively. This includes international tax, transfer pricing, tax strategy, and more. KPMG's consulting services cover various areas such as management consulting, technology advisory, risk management, strategy, and operations. They assist clients in improving performance, driving innovation, and implementing business transformation.

KPMG's deal advisory services encompass mergers and acquisitions (M&A), transaction services, due diligence, valuation, and corporate finance. These services help clients make informed decisions during transactions. KPMG assists clients in identifying and managing risks across various aspects of their business, including cybersecurity, regulatory compliance, data privacy, and internal controls. This includes forensic investigations, dispute resolution, valuation services, and restructuring advisory. KPMG helps clients navigate complex financial challenges and make informed decisions.

KPMG helps organizations leverage technology for growth and efficiency. Their services include IT strategy, digital transformation, cybersecurity, data analytics, and emerging technology adoption.

1.3 Mission and Vision

KPMG's mission statement is: "To turn knowledge and understanding of information, industries, and business trends into value for our firms' clients, our people, and the capital markets." Now for those who may not be aware, KPMG is a B2B service, which is also reflected in their mission statement. The statement clearly lays out that it has its clients' best interests in mind. The use of "Knowledge and Understanding," in the statement alludes to their growing network of highly talented people. It goes without saying that KPMG hires only the best financial experts.

KPMG, in its mission statement, alludes to the fact that they understand every industry, which at first glance may seem outrageous, but it isn't. **"To turn knowledge and understanding of information, industries and business trends into value for our firms' clients…"** merely points to the fact that the company if and when required hires people who specialize in the financial matters of those industries so that they can cater to the needs of their clients. So, in a way, that network of talent we mentioned above contributes to the company's ability to expand its network of clients across all industries.

KPMG's vision statement: **"Distinguished by the learning opportunities offered, through a platform which shares the expertise, knowledge, and experience of professionals across a global network."**

1.4 About the Purpose of KPMG AU Virtual Internship

The KPMG Data Analytics Virtual Internship is designed to help you gain a practical insight into the work we do at KPMG and it is an opportunity for you to build your career skills and experience.

It is not the intention for KPMG to hire students based on their performance in this virtual internship and completing the program is not a pre-requisite when applying for our current vacancies. This virtual internship is an opportunity for you to demonstrate your interest in KPMG and we will consider students favorably who complete the program.

Due to the number of participants potentially completing this program, KPMG will not be able to assess your individual work or provide you with specific feedback. However, after submitting your completed work, you will be able to access model work prepared by our consultants so you can see how someone from KPMG would approach the task.

CHAPTER-2

ABOUT THE DEPARTMENT

2.1 About the Industry Automation Department

It seems like you're interested in learning about the theory related to the field of Industrial Automation. Industrial Automation is a multidisciplinary field that combines various engineering disciplines, computer science, and control systems theory to optimize and control industrial processes automatically. Here are some key theories and concepts associated with the Industrial Automation department.

Control theory is at the core of industrial automation. It deals with designing and analyzing control systems that regulate the behavior of dynamic systems. This theory involves concepts like feedback control, stability, controllability, observability, and PID (Proportional-Integral-Derivative) controllers. In industrial automation, feedback control systems are used to monitor the output of a process and adjust the inputs to maintain the desired output. This involves comparing the actual output with the desired output and making corrective adjustments as needed.

Programmable Logic Controllers (PLCs) are widely used in industrial automation to control machinery and processes. Understanding ladder logic, which is a graphical programming language used in PLCs, is essential for designing and implementing automated control logic. Process modeling involves creating mathematical representations of industrial processes. These models help engineers understand the behavior of processes and design effective control strategies.

2.2 About the Skill Development Department

It appears that you're interested in learning about the Skill Development department. Skill development is a critical aspect of education and workforce preparation, aiming to equip individuals with the knowledge, abilities, and practical expertise needed to excel in various professional fields. The Skill Development department typically focuses on providing training, education, and resources to enhance specific skills that are in demand across industries. Here's an overview of what the Skill Development department entails.

The department often starts by assessing the existing skills of individuals. This assessment helps identify areas where improvement is needed and assists in tailoring training programs to suit individual needs. Skill development involves designing comprehensive and structured curricula that cover the theoretical foundations and practical applications of the targeted skills. The curricula are often designed to align with industry standards and requirements. The department develops and conducts training programs that can be short-term workshops, online courses, certification programs, or longer-term vocational courses. These programs aim to bridge the gap between the skills possessed by individuals and the skills demanded by the job market.

2.2 Main Areas

Mainly focuses on providing hands-on experience to students of graduation and postgraduation cadre. The company's motive is to bridge the gap between student community and current technologies by conducting workshops on recent technological advancements like that of PCB, ARM, Arduino, Robotics, Copters and Android tools.

The training offered to interns here mainly focuses on the R&D (Research and Development) side of the product design taken up by the company. The training offered by KPMG during internship takes place on 2 different levels. In the first level of training, the facilitators take you through the basics of product design and fundamentals of Research and Development(R&D).

2.3 Responsibilities

Able to offer one – stop service with small, medium to mass production with competitive engineering, quality and OTD assurance that customized all your request. Also, the company offers their customers best pricing in the market. Capable to deliver / export the products across the country. Our staff are capable of exploring updated technologies in the market to provide our customers to their needs / requirements. KPMG also provides Reverse engineering services on PCBs to our clients.

CHAPTER-3

DATA ANALYTICS

3.1 Exploring Insights Through Data

In the modern era, where data is being generated at an unprecedented pace, the ability to harness and derive valuable insights from it has become a critical skill. The realm of data analytics plays a pivotal role in uncovering patterns, trends, and meaningful information from vast and complex datasets. This report encapsulates a transformative journey undertaken during an internship in the field of data analytics, where real-world data was transformed into actionable knowledge.

Over the course of this internship, the dynamic landscape of data analytics was navigated, revealing its applications across diverse industries and scenarios. This document not only outlines the methodologies and tools employed to analyze data but also highlights the practical challenges faced and the innovative solutions developed to overcome them. By delving into the intricacies of data collection, preprocessing, analysis, and visualization, this report provides a comprehensive account of the experiences, lessons, and achievements gained during the internship.

The importance of data-driven decision-making cannot be overstated, and this internship provided an immersive opportunity to witness this firsthand. From formulating hypotheses to conducting rigorous analyses, every step in the process was a learning experience that underscored the significance of critical thinking, attention to detail, and the ability to draw meaningful conclusions from data noise. The insights gleaned not only contributed to the host organization's goals but also expanded the horizons of knowledge for the intern.

Throughout this report, the focus remains on transparency and clarity, mirroring the very principles that govern effective data analytics. Each project undertaken during the internship is dissected, elucidating the methodologies adopted, the challenges surmounted, and the outcomes achieved. By embracing both successes and setbacks, this report provides a candid representation of the growth curve experienced in the realm of data analytics.

As the world continues to embrace the digital age, the demand for professionals skilled in data analytics continues to rise. This report serves as a testament to the value of practical experience in honing these skills, showcasing not only technical proficiency but also the ability to adapt, innovate, and collaborate in a fast-paced and ever-evolving field.

In the following sections, the projects, methodologies, and insights gained during the internship are meticulously detailed, providing a comprehensive account of the journey into the world of data analytics.

3.2 Definition of Data Analytics

Data analytics is the systematic process of examining raw data to derive meaningful insights, identify patterns, discover correlations, and make informed decisions. It involves using various techniques, tools, and methodologies to transform data into valuable knowledge. Data analytics encompasses a wide range of activities, from data collection and preprocessing to exploratory analysis and generating actionable insights. The primary goal of data analytics is to uncover hidden trends and relationships within data, enabling organizations to optimize operations, improve strategies, and gain a competitive edge.

3.3 Importance of extracting value from data

In today's data-driven world, organizations collect vast amounts of data from various sources. However, raw data alone has limited utility. Extracting value from data through analytics is crucial for several reasons:

3.3.1 Informed Decision-Making

Analyzing data allows organizations to make informed decisions based on evidence rather than intuition or guesswork. Insights drawn from data analytics guide strategic choices across departments and industries.

3.3.2 Competitive Advantage

Companies that effectively leverage data analytics gain a competitive edge by identifying market trends, understanding customer behavior, and optimizing processes.

3.3.3 Operational Efficiency

Analyzing data helps identify inefficiencies and bottlenecks in operations, leading to streamlined processes and cost savings.

3.3.4 Personalization

Data analytics enables businesses to understand individual customer preferences and tailor offerings to meet their needs, resulting in improved customer experiences.

3.3.5 Risk Management

Analyzing historical data helps assess risks and vulnerabilities, allowing organizations to make informed risk management decisions.

3.3.6 Innovation

Data analytics uncovers new opportunities and potential areas for innovation by revealing patterns that were previously hidden.

3.3.7 Resource Optimization

Data analytics helps allocate resources effectively by identifying areas of overuse or underutilization.

3.3.8 Continuous Improvement

By analyzing data, organizations can track their performance over time, identify areas for improvement, and implement changes to enhance outcomes.

3.3.9 Scientific Approach

Data analytics brings a scientific approach to decision-making, relying on evidence rather than assumptions.

3.3.10 Strategic Planning

Analytics provides insights that inform long-term strategic planning, helping organizations align their goals with data-driven insights.

3.4 Exploring the Power of Data Quality and Presentation

Data quality refers to the accuracy, completeness, reliability, and consistency of the data being used for analysis. It's crucial to ensure that the data used in analytics is trustworthy and free from errors or inconsistencies. The concept highlights the significance of clean and reliable data as the foundation for generating meaningful insights.

Data presentation involves effectively communicating insights and findings derived from data analysis to various stakeholders. It's not just about generating insights; it's about making those insights understandable and actionable. The concept emphasizes the art of storytelling with data and using visuals to convey complex information.

Importance of extracting value from data in today's data-driven world, organizations collect vast amounts of data from various sources. However, raw data alone has limited utility.

3.5 Extracting value from data through analytics is crucial for several reasons

Data analytics is the systematic process of examining raw data to derive meaningful insights, identify patterns, discover correlations, and make informed decisions. It involves using various techniques, tools, and methodologies to transform data into valuable knowledge. Data analytics encompasses a wide range of activities, from data collection and preprocessing to exploratory analysis and generating actionable insights. The primary goal of data analytics is to uncover hidden trends and relationships within data, enabling organizations to optimize operations, improve strategies, and gain a competitive edge.

3.5.1 Data Collection

Data collection is the process of gathering raw data from various sources. This can include structured data from databases, spreadsheets, surveys, as well as unstructured data like text, images, and videos. The quality and relevance of the collected data directly impact the accuracy of the insights derived from analysis. Proper data collection involves defining the scope, identifying data sources, and designing data collection methods such as surveys, sensors, or web scraping.

3.5.2 Data Cleaning and Preprocessing

Data cleaning and preprocessing involve transforming raw data into a usable format for analysis. It includes tasks like handling missing values, removing duplicates, correcting inconsistencies, and standardizing data. This step is crucial as real-world data is often messy and incomplete. By cleaning and preprocessing the data, analysts ensure that the subsequent analysis is accurate and reliable.

3.5.3 Data Exploration and Analysis

Data exploration involves investigating the dataset to understand its structure, patterns, and characteristics. It includes tasks such as creating summary statistics, visualizations, and identifying trends and outliers. Exploratory data analysis helps analysts gain insights into the dataset's nature before applying more advanced analysis techniques. It provides a foundation for hypothesis generation and guides subsequent steps in the analysis process.

3.5.4 Insights Generation

Insights generation is the heart of data analytics. It involves using various statistical, machine learning, and data mining techniques to extract meaningful patterns and information from the data. Analysts create models, perform calculations, and run algorithms to uncover relationships, correlations, and trends. The goal is to transform raw data into actionable insights that inform decision-making. These insights can include predictions, recommendations, or a deeper understanding of a situation.

3.6 Types of Data Analytics

3.6.1 Descriptive Analytics

Descriptive analytics involves examining historical data to understand what has happened in the past. It focuses on summarizing and visualizing data to gain insights into patterns, trends, and key metrics. Descriptive analytics helps answer questions like "What happened?" and provides a foundation for understanding a situation. Example: Analyzing sales data over the past year to identify which products had the highest sales volume during specific months.

3.6.2 Diagnostic Analytics

Diagnostic analytics goes beyond describing what happened and delves into why it happened. It involves exploring relationships and dependencies within the data to uncover the underlying causes of certain outcomes. Diagnostic analytics aims to provide insights into the factors influencing past events Example: Investigating why website traffic dropped in a particular week by analyzing marketing campaigns, website changes, and external events that might have impacted user behavior.

3.6.3 Predictive Analytics

Predictive analytics involves using historical data and statistical algorithms to make predictions about future outcomes. It leverages patterns and trends identified in past data to forecast what might happen next. Predictive analytics is focused on answering the question "What is likely to happen?" Example: Using customer purchasing behavior and demographic data to predict which products a customer is likely to buy next.

3.6.4 Prescriptive Analytics

Prescriptive analytics takes predictive analysis a step further by recommending actions to optimize or address predicted outcomes. It combines data, algorithms, and

business rules to suggest the best course of action to achieve desired goals or avoid potential issues. Prescriptive analytics provides actionable insights and answers the question "What should we do?"

Example: Recommending specific pricing strategies based on predicted market trends and competitor activities to maximize revenue.

3.7 Importance of Data Quality in Analytics

Data quality refers to the accuracy, completeness, reliability, and relevance of the data used in analytics processes. It's crucial to recognize that the accuracy and reliability of the insights derived from data analytics depend heavily on the quality of the data being analyzed. Here's why data quality is of utmost importance in analytics:

3.7.1 Informed Decision-Making

Accurate and high-quality data is the foundation of making informed and reliable decisions. Without trustworthy data, the conclusions drawn from analytics can be misleading and may lead to incorrect actions.

3.7.2 Building Trust

Organizations and stakeholders need to trust the insights generated from analytics. Consistently accurate and reliable data fosters trust in the analytics process and the decisions made based on those insights.

3.7.3 Effective Strategy Formulation

Business strategies are often informed by analytics. If the underlying data is inaccurate, the strategies developed based on flawed insights could lead to misallocations of resources and missed opportunities.

3.7.4 Detection of Anomalies

Poor data quality can mask or distort anomalies and outliers. Identifying these anomalies can lead to valuable insights, especially in fraud detection, quality control, and risk assessment.

3.7.5 Long-Term Planning

Historical data is crucial for predictive analytics and future planning. If historical data is inaccurate or incomplete, predictions and long-term planning may be flawed.

3.7.6 Minimizing Costs

Poor data quality can lead to unnecessary costs. For example, marketing efforts targeting incorrect customer segments can result in wasted resources.

3.8 Impact of Poor Data Quality on Insights

Poor data quality can significantly compromise the accuracy and reliability of the insights generated through data analytics. Here are some key impacts of poor data quality:

3.8.1 Inaccurate Conclusions

Analytics performed on incorrect or incomplete data can lead to inaccurate conclusions and incorrect assumptions about trends and patterns.

3.8.2 Biased Insights

Poor data quality can introduce biases into the analysis, leading to skewed results. This can have serious implications, especially in decision-making processes that rely heavily on analytics.

3.8.3 Missed Opportunities

Insights drawn from low-quality data might not reflect actual opportunities in the market or areas for improvement within the organization.

3.8.4 Incorrect Predictions

Predictive models heavily rely on historical data. Poor data quality can lead to models that fail to predict future trends accurately.

3.8.5 Damaged Reputation

Using flawed insights to make decisions can lead to undesirable outcomes, damaging the reputation of the individuals or organizations relying on those insights.

3.8.6 Resource Wastage

Poor data quality can lead to misguided resource allocation. Organizations might invest in initiatives based on incorrect insights, resulting in wasted time, money, and effort.

3.9 Strategies for Maintaining Data Accuracy and Completeness

Data accuracy and completeness refer to the quality of the data used for analysis. Inaccurate or incomplete data can lead to flawed insights and incorrect conclusions, potentially leading to misguided decisions.

3.10 Data Validation Techniques

3.10.1 Purpose of Data Validation

Data validation aims to confirm whether the data entered or imported into a system or database is accurate, complete, and conforms to predefined standards. It prevents erroneous or inappropriate data from corrupting databases, analytics, and reports.

3.10.2 Common Data Validation Techniques

Format Validation: Ensuring that data follows a specific format. For example, validating that phone numbers are in the correct format.

3.10.3 Range Validation

Verifying that data falls within acceptable numerical or categorical ranges. This is often used for quantities, ages, grades, etc.

3.10.4 Presence Check

Confirming that essential fields are not left blank or null. This prevents missing critical data.

3.11 Data Validation Workflow

3.11.1 Data Collection

Gather data from various sources, such as manual entry, forms, or automated data feeds.

3.11.2 Preprocessing

Clean and transform data to ensure uniformity and consistency.

3.11.3 Validation Rules Definition

Define the rules or criteria that data must adhere to for validation.

3.11.4 Validation Process

Apply the defined rules to the data, identifying any discrepancies or errors.

3.11.5 Error Handling

Address identified errors by correcting or removing invalid data.

3.12 Recognizing anomalies and Outliers

Recognizing anomalies and outliers is a crucial aspect of data analytics that involves identifying data points that deviate significantly from the expected patterns or the norm within a dataset. Anomalies are data points that exhibit behaviors different from the majority, while outliers are extreme values that fall far outside the typical range of values.

3.12.1 Anomalies

Anomalies are data points that don't conform to the general trends or patterns of the rest of the dataset. They can be caused by various factors such as errors in data collection, data corruption, or rare events. Recognizing anomalies is important because they might represent significant occurrences that require further investigation. For instance, in a credit card transaction dataset, an anomaly could be an unusually large purchase that might indicate fraudulent activity. In industrial equipment monitoring, an anomaly might represent a malfunction or breakdown in machinery.

3.12.2 Outliers

Outliers are data points that are significantly different from other values in the dataset, usually by being much higher or lower. Outliers can arise due to a variety of reasons, including measurement errors, experimental noise, or genuine extreme events. Identifying outliers is valuable for understanding the distribution and behavior of the data. For example, in a dataset of income levels, an extremely high income compared to the rest could be an outlier that may provide insights into the economic landscape.

CHAPTER-4

TASK PERFORMED

Fig. 3.1 Data Analytics

4.1 Table of Schedule

WEEKS	NATURE OF WORK
Week 1	Data Quality Assessment
Week 2	Data Insights
Week 3	Data Insights and Presentation

WEEK 1

4.2 Data Quality Assessment

Data Quality Assessment refers to the process of evaluating and ensuring the accuracy, reliability, and completeness of data used in a business or analytical context. This assessment is a crucial step to guarantee that the insights and decisions drawn from the data are valid and trustworthy. Here's an explanation of Data Quality Assessment within the framework of the KPMG Virtual Internship.

In the context of the KPMG Virtual Internship, the Data Quality Assessment refers to the process of evaluating and ensuring the accuracy, reliability, and completeness of data used in a business or analytical context. This assessment is a crucial step to guarantee that the insights and decisions drawn from the data are valid and trustworthy. Here's an explanation of Data Quality Assessment within the framework of the KPMG Virtual Internship.

4.2.1 Data Collection and Understanding

The process begins by understanding the data being used, including its sources, formats, and attributes. This involves gaining insights into the purpose of the data, how it's collected, and the context in which it will be utilized.

4.2.2 Data Profiling

Data profiling involves analyzing the data to identify its structure, patterns, and potential issues. This step helps in uncovering anomalies, missing values, duplicates, and other irregularities that might affect data quality.

4.2.3 Data Cleaning

After identifying data issues, the next step is data cleansing. This involves correcting errors, filling in missing values, and removing duplicates. The goal is to bring the data to a consistent and accurate state.

4.2.4 Data Accuracy

Accuracy refers to how closely the data reflects the real-world entities or events it represents. During the Data Quality Assessment, checks are performed to ensure that the data accurately reflects the facts it is meant to represent.

4.2.5 Data Completeness

Completeness refers to whether all the required data is present. Incomplete data can lead to skewed or incorrect insights. The assessment involves checking for missing values and determining if they can be filled or if additional data needs to be collected. Data Consistency ensures that the data remains uniform and coherent across different sources and time periods. Inconsistencies can lead to confusion and incorrect analysis. The assessment involves identifying and resolving discrepancies.

4.2.6 Data Reliability

Reliability assesses the trustworthiness of the data source and the methods used for data collection. The assessment ensures that data is collected from credible sources using valid methodologies.

4.2.7 Data Timeliness

Timeliness refers to whether the data is up-to-date and relevant for the analysis. Outdated data can lead to inaccurate conclusions. The assessment involves checking data timestamps and considering its relevance to the analysis timeframe.

4.2.8 Data Relevance

Relevance checks whether the data being used is suitable for the specific analysis or decision-making process. Irrelevant data can lead to incorrect insights. The assessment involves evaluating whether the data aligns with the objectives of the analysis.

In the KPMG Virtual Internship, engaging in Data Quality Assessment would likely involve working with simulated data sets, identifying issues, and making necessary corrections to ensure the accuracy and reliability of the data for analysis and reporting purposes. It provides participants with practical experience in handling real-world data challenges, which is essential in various roles within the field of auditing, consulting, and data analytics.

4.3 The background information of task

Sprocket Central Pty Ltd. A medium size bikes & cycling accessories organization, has approached Tony Smith (Partner) in KPMG's Lighthouse & Innovation Team. Sprocket Central Pty Ltd is keen to learn more about KPMG's expertise in its Analytics, Information & Modelling team.

Smith discusses KPMG's expertise in this space (you can read more here). In particular, he speaks about how the team can effectively analyze the datasets to help Sprocket Central Pty Ltd grow its business. Primarily, Sprocket Central Pty Ltd needs help with its customer and transactions data. The organization has a large dataset relating to its customers, but their team is unsure how to effectively analyze it to help optimize its marketing strategy.

4.4 Data Quality Assessment Code on Jupyter Notebook

kpmg1-4-1

August 30, 2023

```
[4]: from google.colab import files
     files.upload()
```

```
[6]: import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     %matplotlibinline
```

UsageError: Line magic function `%matplotlibinline` not found.

```
[7]: train = pd.read_excel('sh(1).xlsx')
     train.head()
```
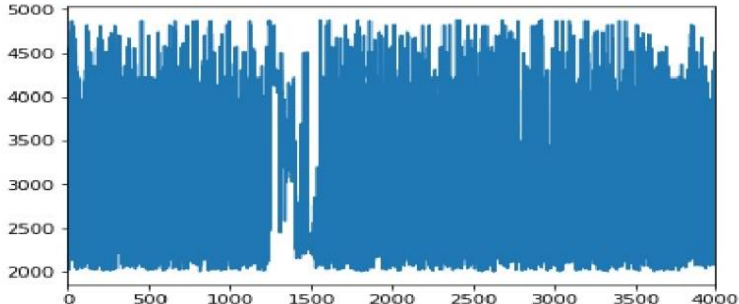
```
[103]: train.isnull().sum()
```

```
[103]: customer_id          0
       address              0
       postcode             0
       state                0
       country              0
       property_valuation   0
       dtype: int64
```

Fig.4.4.1 Importing the File

```
[104]: train['postcode'].plot()
```

[104]: <matplotlib.axes._subplots.AxesSubplot at 0x7f67a3976240>


```
[105]: train['country'].value_counts()
```

```
[105]: Australia    3999
       Name: country, dtype: int64
```

```
[106]: train['property_valuation'].plot()
```

[106]: <matplotlib.axes._subplots.AxesSubplot at 0x7f67a0f8b9b0>
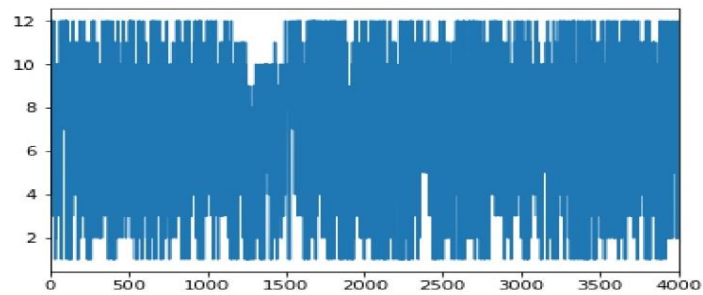
```
[107]:  submission = pd.DataFrame({
            "customer_id": train["customer_id"],
            "address": train["address"],
            "postcode": train["postcode"],
            "state": train["state"],
            "country": train["country"],
            "property_valuation": train["property_valuation"]
        })
        submission.to_csv('./submission.csv', index=False)
        print(submission.head(10))

            customer_id               address  ...   country property_valuation
        0             1     060 Morning Avenue  ...  Australia                 10
        1             2     6 Meadow Vale Court  ... Australia                 10
        2             4       0 Holy Cross Court  ... Australia                  9
        3             5      17979 Del Mar Point  ... Australia                  4
        4             6          9 Oakridge Court  ... Australia                  9
        5             7         4 Delaware Trail  ...  Australia                  9
        6             8     49 Londonderry Lane  ...  Australia                  4
        7             9          97736 7th Trail  ...  Australia                 12
        8            11      93405 Ludington Park  ... Australia                  8
        9            12   44339 Golden Leaf Alley  ... Australia                  4

        [10 rows x 6 columns]
```

Fig.4.4.2 Data Visualization

```
[0]:  from google.colab import files
      files.download('submission.csv')

[0]:
```

Fig.4.4.3 Converting data into CSV file

WEEK 2

4.5 Data Insights

Data insights are generated through the application of data analysis techniques, which can include statistical analysis, data visualization, machine learning, and more. These techniques help in discovering patterns, trends, correlations, and anomalies within the data.

4.5.1 Exploratory Data Analysis (EDA)

EDA involves initially exploring the data to identify its characteristics and patterns. This step might involve creating summary statistics, histograms, scatter plots, and other visualizations to gain an understanding of the data's distribution and relationships.

4.5.2 Descriptive Analytics

Descriptive analytics involves summarizing and describing historical data to provide context and understanding. It answers questions like "What happened?" and often includes metrics like averages, counts, and percentages.

4.5.3 Diagnostic Analytics

Diagnostic analytics aims to understand the reasons behind certain events or trends. It involves digging deeper into the data to uncover the factors that led to specific outcomes.

4.5.4 Predictive Analytics

Predictive analytics uses historical data to make predictions about future events or trends. Machine learning algorithms are often used to build predictive models that can forecast outcomes based on patterns in the data.

4.5.5 Prescriptive Analytics

Prescriptive analytics goes beyond prediction to suggest actions that can be taken to achieve desired outcomes. It provides recommendations based on the analysis of data and possible scenarios.

4.5.6 Data Visualization

Visualizing data through charts, graphs, and dashboards makes it easier to understand and communicate insights. Effective data visualization enhances the ability to spot trends and patterns quickly.

4.5.7 Business Context

Data insights are valuable when they are aligned with the business's goals and objectives. Understanding the business context helps in identifying which insights are most relevant and actionable.

4.5.8 Hypothesis Testing

In data analysis, hypotheses are formulated and tested to validate assumptions or answer specific questions. This involves setting up statistical tests to determine if observed patterns are statistically significant.

4.5.9 Decision-Making

The ultimate goal of deriving data insights is to inform decision-making. Insights guide organizations in making informed choices that can lead to improved processes, increased efficiency, and better outcomes.

In the KPMG Virtual Internship, engaging with data insights might involve tasks such as analyzing financial data to identify trends, assessing customer behavior patterns, or making recommendations based on the results of the analysis. It provides participants with hands-on experience in using data to generate valuable insights that drive informed decision-making, a critical skill in the fields of auditing, consulting, and data analytics.

4.6 Data Insights Code on Jupyter Notebook

newcustomerntb

August 30, 2023

```
[1]: #Get all packages needed
     import pandas as pd
     import numpy as np
     import datetime
     from matplotlib.ticker import FuncFormatter
     import matplotlib.pyplot as plt
     import warnings
     import math as ma
     warnings.filterwarnings("ignore")
```

```
[2]: kmpgex = pd.ExcelFile('KPMG_final.xlsx')
     kmpgex.sheet_names
     df = pd.read_excel('KPMG.xlsx', sheetname="NewCustomerList")
     #list(df)
```

```
[3]: # rename for easier analysis
     df.rename(columns={"Note: The data and information in this document is␣
      ␣reflective of a hypothetical situation and client. This document is to be␣
      ␣used for KPMG Virtual Internship purposes only. ":"fname"}, inplace = True)
     df.rename(columns={"Unnamed: 1":"lname",
                        "Unnamed: 2":"gender",
                        "Unnamed: 3":"3y_bike_purchases",
                        "Unnamed: 4":"DOB",
                        "Unnamed: 5":"JT",
                        "Unnamed: 6":"Category"}, inplace = True)
     df.rename(columns={"Unnamed: 7":"wealth_segement",
                        "Unnamed: 8":"D_Indicator",
                        "Unnamed: 9":"owns_car",
                        "Unnamed: 10":"tencure",
                        "Unnamed: 11":"address",
                        "Unnamed: 12":"postcode"}, inplace = True)
     df.rename(columns={"Unnamed: 13":"state",
                        "Unnamed: 14":"country",
                        "Unnamed: 15":"prop_val",
                        "Unnamed: 21":"rank",
                        "Unnamed: 22":"value"}, inplace = True)
     df=df.iloc[1:]
```

Fig.4.6.1 Importing files

```
df
```

```
[3]:         fname         lname  gender 3y_bike_purchases                DOB  \
     1       Chickie       Brister    Male                86         1957-07-12
     2         Morly        Genery    Male                69         1970-03-22
     3       Ardelis      Forrester  Female               10  1974-08-28 00:00:00
     4        Lucine         Stutt  Female               64         1979-01-28
     5       Melinda        Hadlee  Female               34         1965-09-21
     6         Druci       Brandli  Female               39         1951-04-29
     7      Rutledge         Hallt    Male                23         1976-10-06
     8        Nancie          Vian  Female               74         1972-12-27
     9          Duff      Karlowicz   Male                50         1972-04-28
     10      Barthel        Docket    Male                72         1985-08-02
     11     Rockwell        Matson    Male                94         1995-01-01
     12      Wheeler        Winward    Male               48         1999-08-30
     13         Olag           NaN    Male                60         1990-05-13
     14        Melba       Spellacy  Female               38         1976-12-09
     15       Mandie        Feares  Female               32         1964-04-19
     16        Dukie         Swire    Male                88         1954-03-31
     17     Marcelia     Monkleigh  Female               61         1993-08-22
     18    Winnifred  Beswetherick  Female               83  1976-06-08 00:00:00
     19       Odilia         Quick  Female               65         1938-11-09
     20        Karly      Willavize  Female                2         1954-08-12
     21       Teddie       Burchill    Male               11         1968-12-21
     22       Gaston      Dallaghan    Male               44         1993-09-29
     23         Otis         Ottey    Male                26         1998-02-05
     24     Tabbatha       Averill  Female                5  1977-12-17 00:00:00
     25        Brena      Schnitter  Female               78         1982-10-11
     26       Rourke       Gillbard    Male               11         1945-08-03
     27        Dyane       Burwell  Female               19         1952-06-27
     28     Claudine       Barstowk  Female               71         1966-07-20
     29      Blinnie          Roze  Female               84         1960-12-10
     30        Rhona     De Freyne  Female               45         1960-11-22
     ...         ...           ...     ...              ...                ...
     971        Mavra         Finan  Female               30         1967-08-23
     972        Frieda       Tavinor  Female               43         1999-03-04
     973      Ellwood        Budden    Male               82         1998-06-03
     974         Alex      Patshull  Female               37         1966-01-02
     975      Aundrea       Outridge  Female               77         2001-01-24
     976         Amby        Bodega    Male               63         1968-06-12
     977         Esme       Pilipets    Male               15         1967-05-06
     978      Beverly       Domnick  Female               56         1938-12-10
     979      Artemis       Swanson    Male               77         1977-02-12
     980       Daryle      Marginson    Male               93         1986-06-27
     981         Tyne       Anshell  Female               71         1992-04-08
     982        Leona       Shorrock  Female               83         1951-08-23
     983     Bertrando        Carass    Male               45         1956-06-25
```

```
     983      0.6800    0.680000    0.578000    983.0   983      0.41
     984      0.8700    1.087500    0.924375    983.0   983      0.41
     985      0.8500    1.062500    1.062500    985.0   985     0.408
     986      0.6000    0.750000    0.637500    985.0   985     0.408
     987      0.7875    0.787500    0.669375    987.0   987       0.4
     988      1.0700    1.070000    0.909500    988.0   988    0.3995
     989      0.9625    0.962500    0.818125    988.0   988    0.3995
     990      1.2500    1.562500    1.562500    988.0   988    0.3995
     991      0.6375    0.637500    0.541875    988.0   988    0.3995
     992      0.8625    0.862500    0.733125    988.0   988    0.3995
     993      1.3500    1.350000    1.147500    993.0   993     0.391
     994      1.1500    1.150000    0.977500    994.0   994    0.3825
     995      0.7700    0.962500    0.818125    994.0   994    0.3825
     996      0.7900    0.790000    0.790000    996.0   996     0.374
     997      0.7600    0.950000    0.807500    997.0   997     0.357
     998      1.0625    1.062500    1.062500    997.0   997     0.357
     999      1.3625    1.362500    1.158125    997.0   997     0.357
     1000     0.5875    0.734375    0.624219   1000.0  1000      0.34

     [1000 rows x 23 columns]
```

Fig.4.6.2 Describe the Data

```
        return mydict

    [0, 29, 0, 0, 17, 106, 165, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0] 23
```

```
[7]: check_unique()["gender"] #no F or M, easier to analyze
```

```
[7]: (3, array(['Male', 'Female', 'U'], dtype=object))
```

```python
[25]: #Analyze the Gender vs Bikes bought
      #array(['F', 'Male', 'Female', 'U', 'Femal', 'M'] from KPMG_VI.py
      gender = [0,0,0] #first is Female, second if Male, third is Unknown
      for each in df['gender']:
          if each[0] == "F":
              gender[0] += 1
          elif each[0] == "M":
              gender[1] += 1
          else:
              gender[2] += 1

      print(gender)
      def gf(x, pos):
          #'The two args are the value and gender'
          return int(x)

      bike = [0,0,0]
      df[['3y_bike_purchases']] = df[['3y_bike_purchases']].apply(pd.to_numeric)
      bike[0] += df['3y_bike_purchases'][df['gender'] == 'Female'].sum()
      bike[1] += df['3y_bike_purchases'][df['gender'] == 'Male'].sum()
      bike[2] += df['3y_bike_purchases'][df['gender'] == 'U'].sum()

      grapho = [gender[0], bike[0], gender[1], bike[1], gender[2], bike[2]]
      graphl = [gender[0]*10, bike[0], gender[1]*10, bike[1], gender[2]*10, bike[2]]
      colorr = (0.3,0.1,0.4,0.6)
      #scale up gender by 10 for easier visualization
      formatter = FuncFormatter(gf)
      x = np.arange(6)
      fig, ax = plt.subplots()
      ax.set_ylabel('Number of people')
      ax.yaxis.set_major_formatter(formatter)
      plt.bar(x, graphl, color = colorr)
      for i in range(len(gender*2)):
          plt.text(x = i-0.3, y = graphl[i]+0.1, s = grapho[i], size = 10)

      plt.xticks(x, ('Female', "F's P", 'Male', "M's P", 'Unknown', "U's P"))
```

```
[513, 470, 17]
```

Fig.4.6.3 Data Analysis

```
[25]: ([<matplotlib.axis.XTick at 0x1f24e80a128>,
        <matplotlib.axis.XTick at 0x1f2502ddfd0>,
        <matplotlib.axis.XTick at 0x1f2502e20b8>,
        <matplotlib.axis.XTick at 0x1f250071f28>,
        <matplotlib.axis.XTick at 0x1f2500575f8>],
       <a list of 6 Text xticklabel objects>)
```
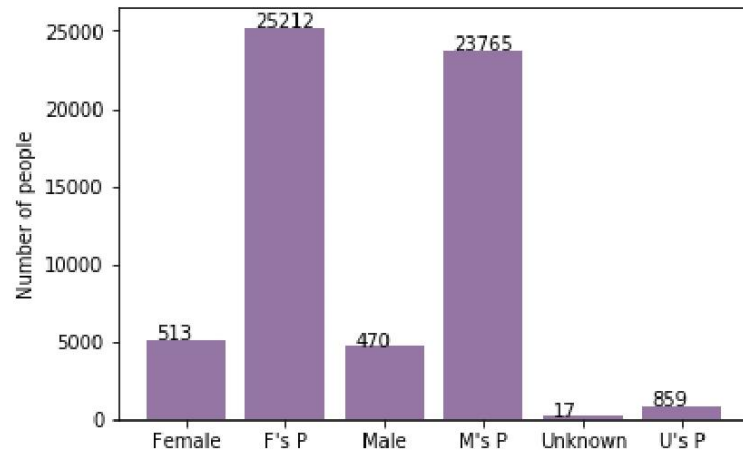


```
[26]: avg  = [0,0,0]
      ss   = sum(bike)
      for i in range(len(avg)):
          avg[i] += round(bike[i]/ss,3)

      x1 = np.arange(3)
      fig1, ax1 = plt.subplots()
      ax1.set_ylabel('Percentage bikes bought by gender')
      plt.bar(x1, avg, color = colorr)
      for i in range(len(avg)):
          plt.text(x = i-0.15, y = avg[i], s = avg[i], size = 10)

      plt.xticks(x1, ('Female', 'Male', 'Unknown'))
```

```
[26]: ([<matplotlib.axis.XTick at 0x1f250557d68>,
        <matplotlib.axis.XTick at 0x1f250447b00>,
        <matplotlib.axis.XTick at 0x1f250583d68>],



       <a list of 3 Text xticklabel objects>)
```


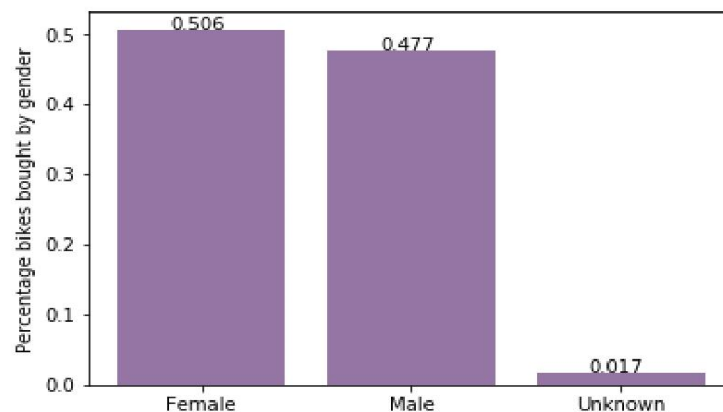**Fig.4.6.4 Data Visualization**

```
[48]: #Analyze on the age vs bikes bought
      #Needs to transform
      df["Age"] = 0
      lenn = len(df["DOB"])
      k      = 0
      for i in range(1, lenn):
          #type is date
          if isinstance(df["DOB"][i], datetime.date):
              tl = len(df["DOB"][i].ctime().split(" "))
              print(int(df["DOB"][i].ctime().split(" ")[tl-1]))
              df["Age"][i] += int(2019 - int(df["DOB"][i].ctime().split(" ")[tl-1]))
          #type is string
          elif isinstance(df["DOB"][i], str):
              tl = len(df["DOB"][i].split("-"))
              df["Age"][i] += int(2019 - int(df["DOB"][i].split("-")[0]))
          #time is float: means it's nan value so do nothing
      #same index as above
      ngenage = [0,0,0]
      j       = 1
      ss      = 0
      stdv    = []
      for each in df['gender']:
          if each[0] == "F" and df['Age'][j] != 0:
              ss   += df['Age'][j]
              stdv.append((df['Age'][j]))
              ngenage[0] += 1
          elif each[0] == "M" and df['Age'][j] != 0:
              ss += df['Age'][j]
              stdv.append((df['Age'][j]))
              ngenage[1] += 1
          elif df['Age'][j] != 0:
              ss += df['Age'][j]
              stdv.append((df['Age'][j]))
              ngenage[2] += 1
          j += 1
      df
```

```
1974
1976
1977
1977
1978
1975
1978
```

```
[49]: print(ss)
      print(ngenage) # as we can see, unknonwn gender will unlikely to have age,␣
       ↪don't care about gender U
      #average age not counting 0 is
      mean_val = round(ss/sum(ngenage),0)
      print(mean_val)
      stdv_val = round(ma.sqrt(1/(sum(ngenage)-1)*sum((stdv - (ss/
       ↪sum(ngenage)))**2)),0)
      print(stdv_val)
```

```
47323
[513, 469, 0]
48.0
17.0
```

```
[54]: #pruchases from age 42 - 34/2, 42, 42 + 34/2
      age_dict = {}
      f1  = []
      bf1 = []
      f2  = []
      bf2 = []
      f3  = []
      bf3 = []
      f4  = []
      bf4 = []
      m1  = []
      bm1 = []
      m2  = []
```

```python
bm2 = []
m3  = []
bm3 = []
m4  = []
bm4 = []
fq = 42-34/2
sq = mean_val
tq = 42+34/2
jjj = 1
for each in df['gender']:
    temp = int(df['Age'][jjj])
    bkt  = int(df['3y_bike_purchases'][jjj])
    if each[0] == "F" and temp != 0:
        if(temp <= fq):
            f1.append(temp)
            bf1.append(bkt)
        elif(fq < temp and temp <= sq):
            f2.append(temp)
            bf2.append(bkt)
        elif(sq < temp and temp <= tq):
            f3.append(temp)
            bf3.append(bkt)
        elif(tq < temp):
            f4.append(temp)
            bf4.append(bkt)
    elif each[0] == "M" and temp != 0:
        if(temp <= fq):
            m1.append(temp)
            bm1.append(bkt)
        elif(fq < temp and temp <= sq):
            m2.append(temp)
            bm2.append(bkt)
        elif(sq < temp and temp <= tq):
            m3.append(temp)
            bm3.append(bkt)
        elif(tq < temp):
            m4.append(temp)
            bm4.append(bkt)
    jjj += 1
dtt = {"Female1":f1,
       "Female2":f2,
       "Female3":f3,
       "Female4":f4,
       "Male1":m1,
       "Male2":m2,
       "Male3":m3,
       "Male4":m4,
```

Fig.4.6.5 Specifing Gender Data
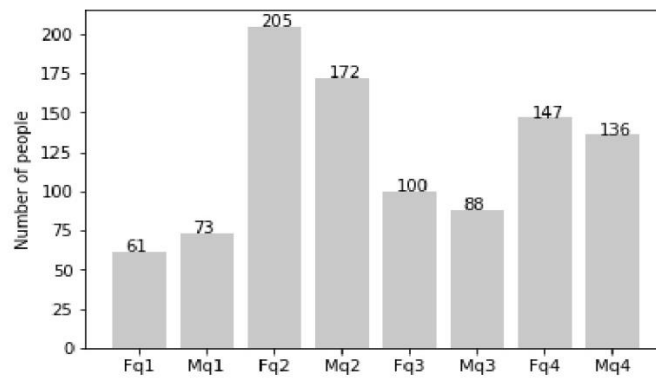
```
[ ]: #We should focus on people around second quadrant (25-48) and fourth quardrant
     > 59
```

```
[58]: #check_unique()["Category"]
      df["Category"].value_counts()
```

```
[58]: Financial Services    203
      Manufacturing         199
      Health                152
      Retail                 78
      Property               64
      IT                     51
      Entertainment          37
      Argiculture            26
      Telecommunications     25
      Name: Category, dtype: int64
```
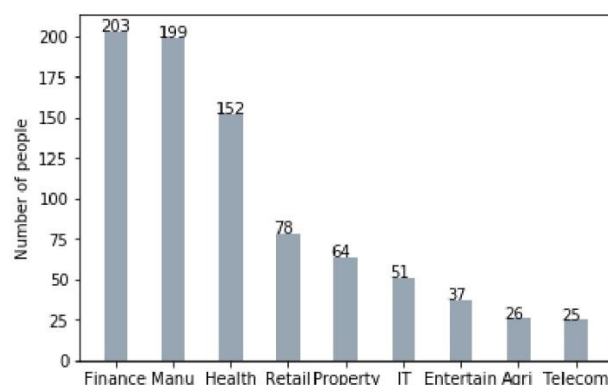
```
[68]: val = [203, 199, 152, 78, 64, 51, 37, 26, 25]
      colorr = (0.2,0.3,0.4,0.5)

      x3 = np.arange(9)
      fig3, ax3 = plt.subplots()
      ax3.set_ylabel('Number of people')
      plt.bar(x3, val, color = colorr, width = 0.4)
      for i in range(len(val)):

          plt.text(x = i-0.25, y = val[i]+0.1, s = val[i], size = 10)

      plt.xticks(x3, ("Finance", "Manu", "Health", "Retail", "Property", "IT",
      "Entertain", "Agri", "Telecom"))
```

```
[68]: ([<matplotlib.axis.XTick at 0x1f24fe20be0>,
        <matplotlib.axis.XTick at 0x1f24fec3c18>,
        <matplotlib.axis.XTick at 0x1f24ff24d30>,
        <matplotlib.axis.XTick at 0x1f24fd83860>,
        <matplotlib.axis.XTick at 0x1f24fd83eb8>,
        <matplotlib.axis.XTick at 0x1f24fd7f588>,
        <matplotlib.axis.XTick at 0x1f24fd7fc18>,
        <matplotlib.axis.XTick at 0x1f24fd612e8>,
        <matplotlib.axis.XTick at 0x1f24fd61978>],
       <a list of 9 Text xticklabel objects>)
```

```
            nsw_nc  += 1
        elif es == "VIC":
            if df["owns_car"][idx] == "Yes":
                vic_car += 1
            else:
                vic_nc  += 1
        elif es == "QLD":
            if df["owns_car"][idx] == "Yes":
                qld_car += 1
            else:
                qld_nc  += 1
        idx += 1
```

```
[89]: graphgen = [nsw_car, nsw_nc, vic_car, vic_nc, qld_car, qld_nc]
      colorr = (0.3,0.5,0.7,0.9)

      x2 = np.arange(6)
      fig2, ax2 = plt.subplots()
      ax2.set_ylabel('Car owns or not in each states')
      plt.bar(x2, graphgen, color = colorr)
      for i in range(len(graphgen)):
          plt.text(x = i-0.2, y = graphgen[i]+0.1, s = graphgen[i], size = 10)

      plt.xticks(x2, ('nsw_car', "nsw_nc", 'vic_car', "vic_nc", 'qld_car', "qld_nc"))
```

```
[89]: ([<matplotlib.axis.XTick at 0x1f24fcd6b38>,
        <matplotlib.axis.XTick at 0x1f2500e7e10>,
        <matplotlib.axis.XTick at 0x1f2500d4c50>,
        <matplotlib.axis.XTick at 0x1f24f764668>,
        <matplotlib.axis.XTick at 0x1f24f764eb8>,
        <matplotlib.axis.XTick at 0x1f24fd4ec18>],
       <a list of 6 Text xticklabel objects>)
```
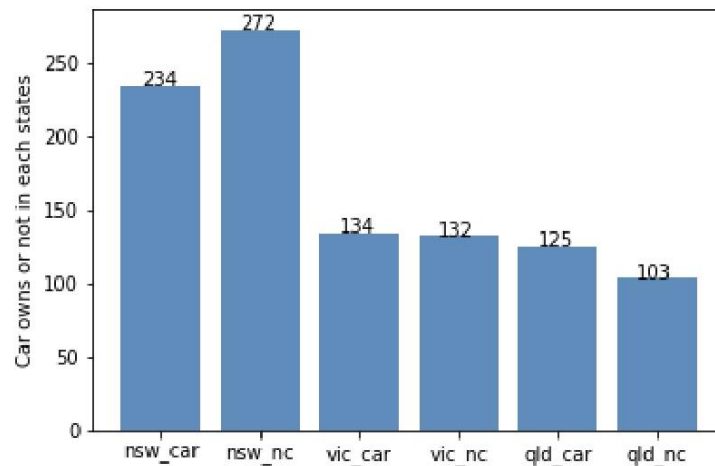

Fig.4.6.6 Insight from the given Data

WEEK 3

4.7 Data Analysis and Exploration

Data insights start with collecting and preparing data from various sources relevant to the business problem or scenario. Exploratory data analysis (EDA) is performed to understand the characteristics of the data, identify patterns, and uncover potential trends or outliers. Statistical analysis, data visualization, and other analytical techniques are used to transform raw data into actionable insights.

4.7.1 Insights Generation

Insights are derived by analyzing the data using statistical methods, machine learning algorithms, and domain expertise. Patterns, trends, correlations, and key findings are identified within the data. Hypotheses might be tested to validate assumptions and draw meaningful conclusions.

4.7.2 Business Relevance

Insights are contextualized within the specific business problem or objective. It's important to ensure that the insights align with the goals of the organization. Understanding the business context helps in selecting which insights are most relevant and valuable.

4.7.3 Insights Interpretation

The insights are interpreted to provide actionable recommendations or solutions to address business challenges. This step involves connecting the dots between the data findings and the potential impact on the business.

4.7.4 Data Visualization

Visualizations such as charts, graphs, dashboards, and heatmaps are used to present insights in an easily understandable and visually appealing manner. Visualizations aid in conveying complex information quickly and effectively.

4.7.5 Presentation Preparation

Based on the insights and visualizations, a presentation is prepared to communicate the findings to stakeholders. The presentation structure is organized logically, with a clear introduction, main points, supporting evidence, and conclusion.

4.7.6 Narrative Creation

The presentation narrative tells a story. It explains the problem, the data analysis approach, the key insights, and the recommended actions. The narrative guides the audience through the insights and helps them understand the significance.

4.7.7 Data-Driven Recommendations

The insights guide the formulation of actionable recommendations that can lead to improved decision-making or operational changes. Recommendations are based on evidence from the data analysis and are tailored to the business context.

4.7.8 Communication Skills

Effective communication skills are essential for presenting insights. Clarity, simplicity, and conciseness are important to ensure that the audience understands the findings.

4.7.9 Stakeholder Engagement

During the presentation, engaging with stakeholders by addressing questions and concerns is crucial for fostering a collaborative environment.

4.7.10 Feedback and Iteration

Feedback from stakeholders may lead to further refinement of insights and recommendations. The presentation may be iterated upon to enhance clarity and alignment with stakeholder needs.

4.7.11 Professionalism

The final presentation reflects professionalism and attention to detail, showcasing a high level of analysis and thoughtful insights.

In the KPMG Virtual Internship, the "Data Insights and Presentation" module likely involves tasks such as analyzing financial data, identifying trends, crafting data-driven recommendations, and preparing a presentation to communicate the findings to a fictional client or team. This experience provides participants with valuable exposure to the process of turning data into actionable insights and effectively conveying those insights through professional presentations.

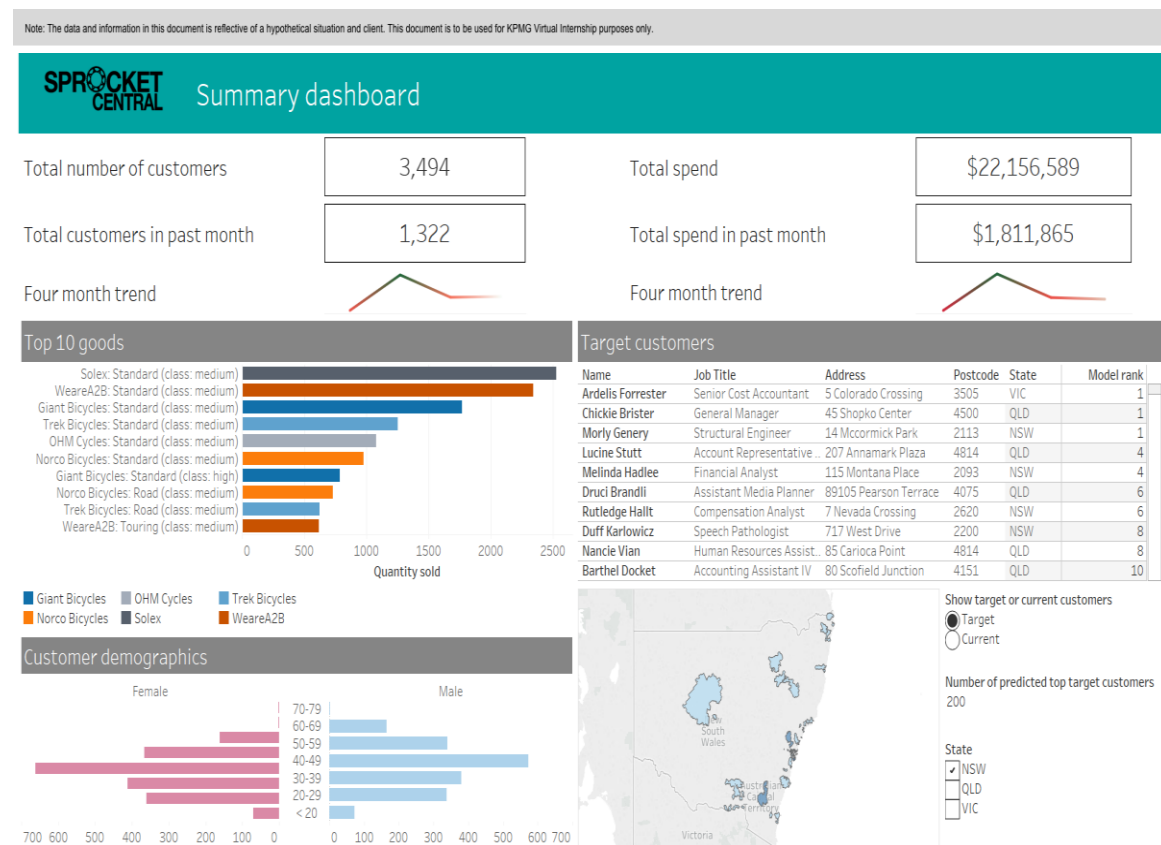## 4.8 Dashboard drawn from Insights

Fig.4.8 Summary Dashboard of Sprocket Central

CHAPTER-5

REFLECTION NOTES

5.1 WORK EXPERIENCE

An internship is a professional introduction to an industry to pursue in the future. Educational counsellors typically arrange internships and try to match you to a company that fits your career goals. Understanding the benefits of doing an internship can help us to decide if this is the right course of action for you.

The following internship training was very helpful to me , I was able to learn new tool Power BI and Excel which will be helpful for us to get good jobs and knowledge. We had to work on building a website using python, firebase as database system. It was a collection of very interactive sessions where we were learning and implementing on the project.

As a beginner to front-end and back-end, I was able to relate things, and grab the new concepts. I am grateful that the training has improved my skills.

As I have gained the hands-on experience in using python for Website development or making. The experience gained here serves as a starting point for my career in programming and can help to build a strong foundation for future work.

5.2 BENEFITS OF INTERNSHIP

5.2.1. Exposure to practical activity

Usually, we students gain mostly theoretical knowledge compare to practical Knowledge, which is not good idea. By doing internship I got to know about how a particular task is done in the company, the overall environment helped me to gain moreover practical knowledge through hands-on. I got to know about how a particular task is done in the company.

5.2.2. Build of strong Resume

This valuable job experience has enabled me to have a clear idea about my career path and the roll which can best suitable for me in cooperating world. Know I am able to mention my skills and experiences in internship honestly in my resume.

5.2.3 Provides work experience

When we enter the workforce for the first time after completing your education or training program. we are probably not going to have much work experience. Internships are the perfect way to fill this gap. We could gain valuable insight into business processes in our chosen field, take part in professional meetings and perform tasks assigned to you in a professional setting.

5.2.4 Builds confidence

When we transition from college or university into full-time employment, we may feel a mix of emotions, including anxiety, excitement, ambition and hope. Internships are a sensible way to make this transition, as we would be learning about the work you aspire to do for the long term. An internship may eliminate the pressure of finding a full-time job quickly and help you to apply the knowledge and skills you have acquired to various practical situations. Internships can act as preliminary employment training, and human resources personnel and managers can observe you in various situations to determine if you might be a valuable addition to the team. It can reduce some of the stress you might otherwise feel if you were looking for a job without internship connections.

5.2.5 TECHNICAL OUTCOME

The outcome of the program is to build a Dashboard, mainly draw Insights, from the given data of Sprocket Central Pty Ltd. Following are interlinked to the python programmer (the logic for the task to be performed). And these are connected to the IDE of Jupyter Notebook, (It is Anaconda based tool). All these are interlinked through Data Analysis.

CHAPTER-6

REFERENCES

6.1 https://www.theforage.com/ This website Which provided the Internship.

6.2https://kpmg.com/in/en/home/careers/students This Company Which provided the Internship.

6.3 Excel Tool which is mostly used.

6.4 Jupyter Notebook is also used and for Dashboard we used Power BI.