

**Data Technician**

|  |
| --- |
|  |

|  |
| --- |
| Name: Alaa Mostafa |
| Course Date: 16/12/24 |
|  |

**Table of contents**

[Day 1: Task 1 2](#_Toc77637984)

[Day 2: Task 1 4](#_Toc1634060488)

[Day 2: Task 2 7](#_Toc152114794)

[Day 2: Task 3 9](#_Toc257844391)

[Day 3: Task 1 10](#_Toc1014152162)

[Day 3: Task 2 13](#_Toc1498274088)

[Dataset: 13](#_Toc1056274673)

[Step 1: Create a Pivot Table 13](#_Toc782776295)

[Step 2: Use the SWITCH Function 13](#_Toc365195726)

[Submission: 14](#_Toc485671904)

[Day 3: Task 3 15](#_Toc1856180793)

[Day 4: Task 1 16](#_Toc381189142)

[Course Notes 19](#_Toc1368242635)

[Additional Information 21](#_Toc305684719)

# Day 1: Task 1

Please complete the below boxes on commons laws and regulations that must be followed when working with customers data, use the below bulleted list to support your answers.

* What is it
* Why is it important
* Provide a real-world example of how you can follow it
* How does it impact working with data
* What could happen if you breached it

|  |  |
| --- | --- |
| Data Protection Act | Data Protection Act is a law that ensures data collected by organisations is used lawfully and securely to protect privacy of individuals. It is important because it protects the sensitive personal data of people that can be easily misused. A good example of applying the law for data security is to provide layers of cyber security for data storages to protect sensitive data, preventing data leaks, and restricting advertisers to use data to make targeted advertisements excessively. The Act impacts working with data by making organisations to collect data lawfully while having some limitations, collecting only the necessary amount, maintaining accuracy, and protecting the data with cyber security methods. Breach of Data Protection Act leads to fines, loss of reputation, operational disruption, and restrictions for an organisation. Individuals can be accused of identity theft or privacy invasion if breaking the law. |
| GDPR | General Data Protection Regulation is a law that makes strict rules for organisations processing personal data. The law requires organisation to make data collection transparent for users, users should be able to demand a record of data that has been collected and to delete it if requested. Similarly to DPA, GDPR also requires to keep data secured, as well as data breaches should be reported. And similarly to DPA, the good real world example of following GDPR is to provide security, minimising data collection, lawful use, but GDPR also provides transparency, retention limits, and allowing users to access the data being collected. Breaching GDPR leads to massive fines, criminal investigation, loss of reputation, and costs to fix security gaps. |
| Freedom of Information Act | Freedom of Information Act is a law that gives people a right to access information that is held by public authorities, it makes sure that users can ask for records and documents, this is important because it helps to build trust between the public and the government while also making the government more transparent. A real world example would be a journalist asking for data about spending on public projects to see if the money is used properly, in this way the law helps to reveal information that otherwise might be hidden. It also has an impact on working with data because people who manage information in public offices need to make sure that it is organised and stored properly, and that personal details are protected before information is shared. Breach of the law may lead to legal issues, damage to trust, and possible penalties for the authority or the person responsible. |
| Computer Misuse Act | Computer Misuse Act was created to restrict individuals and to criminalize illegal access to computer systems without permission, as well as changing data in any way with unauthorised access. Spreading malware such as viruses and worms is also considered a crime. The Act is important because it protects against cyber attacks, and people stealing or sabotaging data. A real life example would be trying to guess another person's password to enter their email account, which is an offence under this act. It impacts working with data in the sense that organisations need to make sure that their systems are secure, employees are taught responsible usage, and any data that has contact with them is protected against computer misuse. In case of the law breach, the penalty could be criminal prosecution, fines, imprisonment, and loss of trust in the individual or the organisation that allowed the abuse to take place. |

# Day 2: Task 1

Please research and complete the following tasks within the retail-sales\_dataset.xlsx document, paste a print screen into the provided boxes below:

1. In the sheet ‘retail\_sales\_dataset’ add all available data between columns A –J into a ‘table’
2. Using the ‘sort’ function, sort ‘Age’ to ‘largest to smallest’
3. Using the ‘SUM’ function, show me the commission total in cell ‘L10’
4. Using the ‘AVERAGE’ function, show me the average commission in cell ‘L11’

|  |  |
| --- | --- |
| Print screen 1 |  |
| Print screen 2 |  |
| Print screen 3 |  |
| Print screen 4 |  |

# Day 2: Task 2

Please research and complete the following tasks within the retail-sales\_dataset.xlsx document, paste print screens into the provided box below:

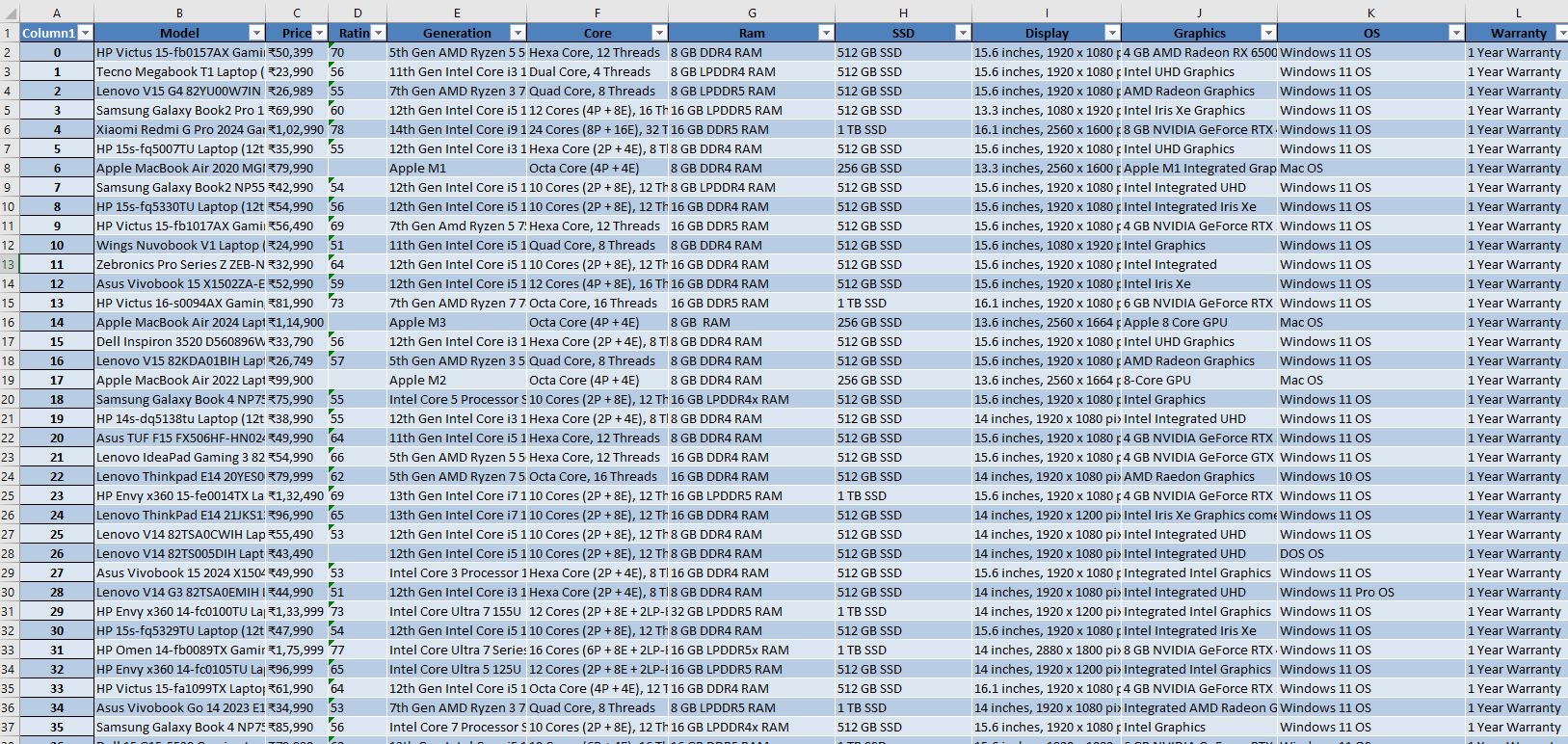


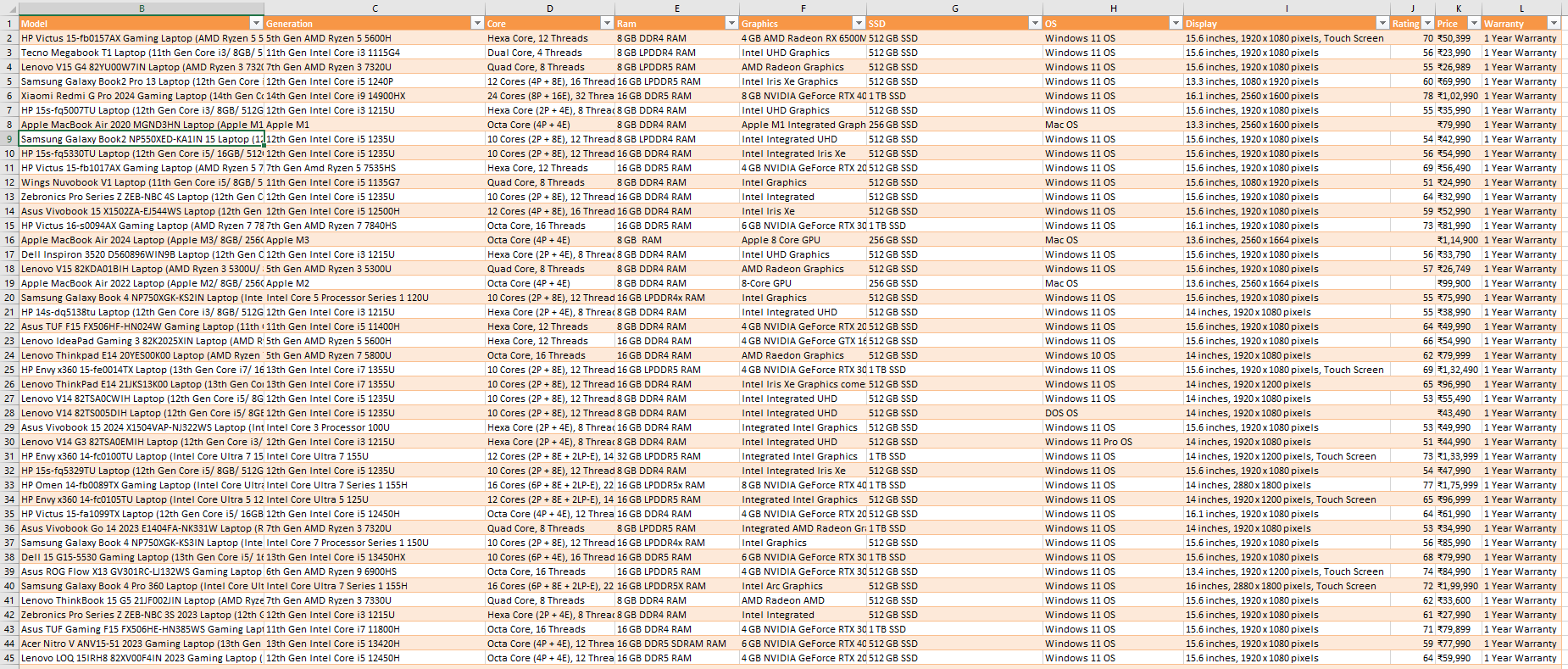
|  |  |
| --- | --- |
| Print screen 1 |  |

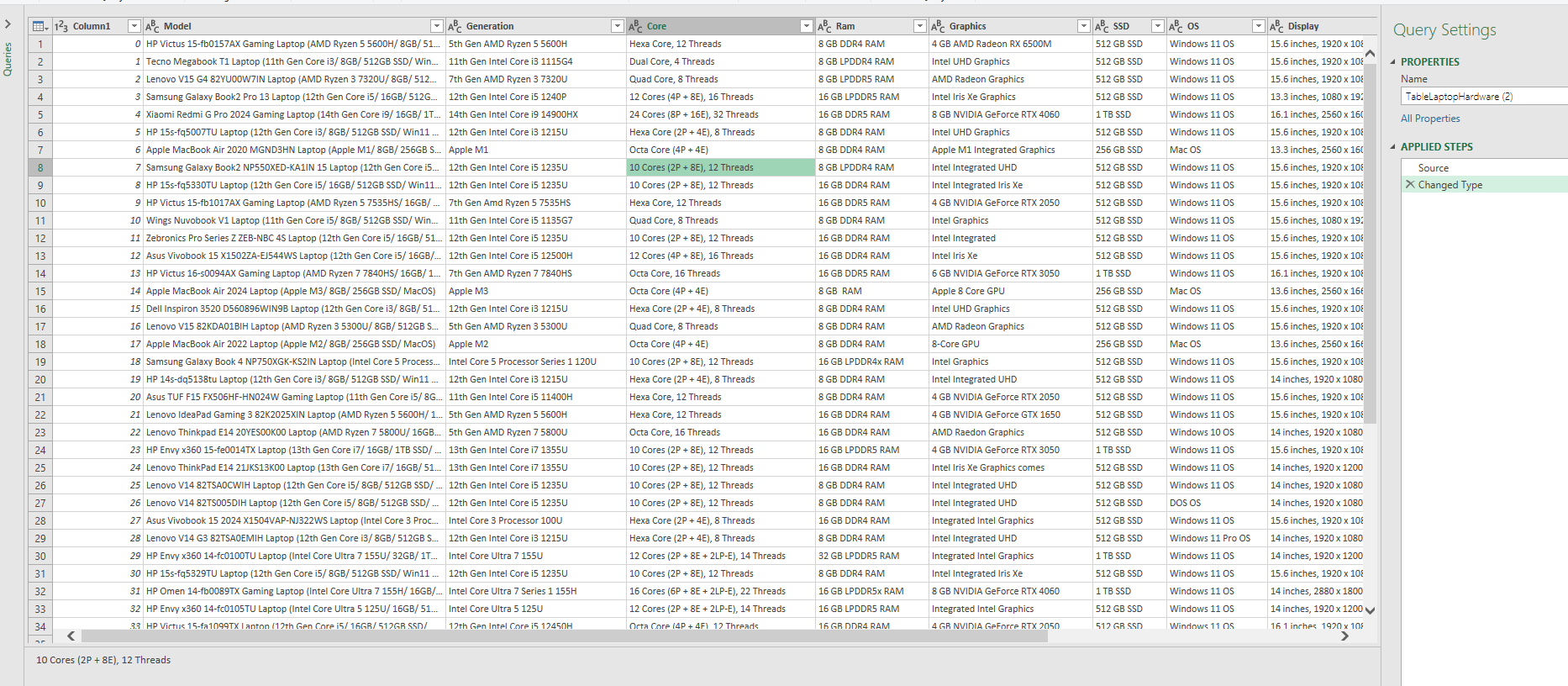
# Day 2: Task 3

Using the skills developed today, have some fun with the data set you have imported. Paste your work below and enjoy!

|  |  |
| --- | --- |
| Print screen 1 |  |







# Day 3: Task 1

Please download the dataset ‘Day\_3\_Task\_1\_Bike\_Sales\_Pivot\_Lab.xlsx’ and the lab instructions.

Do not worry if you do not complete the lab, just working with data and playing with the pivot table will be good experience.

Please paste your final pivot table below and complete the reflection questions:

|  |  |
| --- | --- |
| Print screen 1 |  |
| In which markets does Germany have customers? |  |
| What country has sales in all markets? |  |
| What are the most profitable markets by country, age group, and gender? |  |
| Any other findings? |  |

# Day 3: Task 2

The dataset below tracks the sales performance of different products in various counties in England. Please paste the dataset into a blank Excel workbook. Your task is to:

* **Create a Pivot Table** to summarise the data by county and product.
* **Use the SWITCH function** to categorise products based on their sales volume.

#### **Dataset:**

|  |  |  |
| --- | --- | --- |
| **County** | **Product** | **Sales Volume** |
| Yorkshire | Laptops | 500 |
| Yorkshire | Smartphones | 200 |
| Cornwall | Laptops | 700 |
| Cornwall | Printers | 400 |
| Lancashire | Smartphones | 150 |
| Lancashire | Laptops | 600 |
| Essex | Printers | 800 |
| Essex | Smartphones | 300 |
| Durham | Laptops | 250 |
| Durham | Printers | 300 |
| Greater Manchester | Smartphones | 600 |
| Greater Manchester | Laptops | 400 |

#### **Step 1: Create a Pivot Table**

* Select the dataset (columns A to C).
* Insert a Pivot Table to summarise the data by **County** in the rows and **Products** in the columns. Use **Sales Volume** as the value to be summarised.

#### **Step 2: Use the SWITCH Function**

In a new column next to your data, use the SWITCH function to categorise products based on **Sales Volume** as follows:

* + For sales greater than 600: **"High"**
  + For sales between 300 and 600: **"Medium"**
  + For sales less than 300: **"Low"**

**SWITCH Function Example**:

=SWITCH(TRUE, C2 > 600, "High", C2 >= 300, "Medium", "Low")

* Apply this formula to each row, and check if the products are categorised correctly.

#### **Submission:**

* A completed Pivot Table summarising sales by county and product.
* A new column in the dataset categorising products by sales volume using the SWITCH function.
  + Please paste your completed work below

|  |  |
| --- | --- |
| Print screen 1 |  |

# Day 3: Task 3

Please download the dataset ‘Day\_3\_Task\_3\_Bike\_Sales\_Visualisations\_Lab.xlsx’ and the the lab instructions. Do not worry if you do not complete the lab, just working with data and playing with the charts will be good experience.

Please paste your results below:

|  |  |
| --- | --- |
| Print screen 1 |  |

# Day 4: Task 1

You have been asked to deliver your analysis findings to the board of directors, with your analysis you have identified that customers are leaving your company at the 12-month point, this is typically when they receive their renewal price.

Conduct research and complete the below questions:

|  |  |
| --- | --- |
| How would you prepare for the delivery? | To prepare for the delivery, the identification should be checked to ensure there is no error in the data and data models created – issues within the data model may show misleading information and the presentation may have confusing and unrealistic statements if the data is no correct, as well as visualisations will fail to show realistic scenarios. To present the data analysis, a presentation should be created to provide accessible data and descriptions of the data for the client – a presentation should not include too much text or visualisations so a client will not be overwhelmed by data. The complexity behind the data should be simplified to make it understandable in a common sense context, to make the information accessible for those who are not understanding the whole scene behind the data. Notes are a useful tool to remind the topic of each stage of a presentation, the notes should not be long, in fact it should be practised to remember the topics and to recall the main points that are to be delivered for a client. The presentation is useful to have some keywords or small phrases to remind the main topics. |
| What tools would you use for the delivery? | Tools like Pivot tables, visualisations like graphs, a sequenced presentation, and notes should be used to describe and present the data for a client – visualisations will make it easier for a client to see co-relations and dependencies in the statistics to make decisions and opinion about processes within a company to make further decisions. Presentations are to show data easily and to keep attention with the visualisations and the key topics of the presentation while verbally describing the data and visualisations of data. Aforementioned notes are to be used as a reminder of the key details and things to present and to be described about the data. Pivot tables are also great to sort data, choosing the specific arrays of data and their relation to other arrays to show statistics in a way to see co-relations easier, while making creation of graphs easier too to show them to a client. To find blank places in the data worksheet and to find irrelevant duplicates, it is useful to use conditional formatting, a function COUNTBLANK, and the option to remove duplicates will help to define duplicates and blank data spaces as well as deleting them. |
| What is prospecting and why would you complete this before your delivery? | This should be completed in order to deliver information clearly and precisely while keeping a client’s attention to the details about the data analysis that has been done to record processes within the business. |
| Tell me best practices for public speaking and providing updates to senior leaders | Several of the best practices to keep the audiences attention is to keep voice loud to make everyone hear you and to keep attention, as well as keeping eye contact with the audience to make them feel connected, engaged, it is good to keep the eye contact for 1-2 minutes on a person or several people and changing eye contact to a next group to keep everyone engaged. Looking on a floor or ceiling is a bad practice as it might look strange for the audience and they won’t be engaged as they expect you to talk with them. Describing too much details will overwhelm audience with information, that is why short and simple descriptions are useful to keep presentation within a good time length. Moving around in the room too much and making too much of unnecessary movements like fidgeting hands are not effective and will |
| What will you show the board in your delivery? | The board will show short descriptions and visualisations of data collected and analysed. One slide should only have maximum of 2 or 3 graphs to not overwhelm audience as well as not much words should be written to be short and to deliver information with simplicity. Main data analysis results will be displayed according to the requirements of the client to display the insights that has been asked. |
| How will you articulate the changes that are needed? | To articulate the needed changes, it will be required to create predictive data analysis with use of predictive visualisations to show probable scenarios, what-if scenarios are also a good option to display changes in certain precise scenarios. |
| Provide a list of online resources and videos that will support your preparation for public speaking |  |
| Evaluate tools that provide visualisation.  Tell me what they are.  Tell me what you would choose when delivering your presentation and why | The tools such as Pivot table and graphs are the ones that will be used in a presentation, pivot table is a tool that allows to choose certain data to be displayed in a table that is created automatically, and where rows and columns can be arranged to display certain precise information without any other unnecessary data, it also allows to make quick formula calculations to calculate average, max, min, and sum of a value displayed on a pivot table. It is very flexible to sort and filter the data as well as manipulating data to be displayed, however, it lacks readability even while being a smaller version of a filled data worksheet. Visualisations such as scatter plot, line graphs, and pie charts are the tools to effectively visualise data. Scatter plot is suitable to represent connections between two variables where the position of points gives a clear way to understand correlation or patterns in data. Line graphs are useful to show changes of values over time, they can be easy to read and make it possible to see trends and compare data sets in a simple way. Pie charts are used to represent proportions of categories within a whole, they are effective to show how a dataset is divided into parts, but can be less accurate if too many categories are used. Visualisations overall are very helpful to present findings in a simple way, they make data easier to understand for people who might not be familiar with raw numbers, however, it is important to choose the correct type of graph for the correct purpose, otherwise the data can be misunderstood or confusing. |

|  |
| --- |
| **Course Notes** |

It is recommended to take notes from the course, use the space below to do so, or use the revision guide shared with the class:

|  |
| --- |
|  |

|  |
| --- |
| **Additional Information** |

We have included a range of additional links to further resources and information that you may find useful, these can be found within your revision guide.

**END OF WORKBOOK**

**Please check through your work thoroughly before submitting and update the table of contents if required.**

**Please send your completed work booklet to your trainer.**