

**Data Technician**

|  |
| --- |
|  |

|  |
| --- |
| Name: Saharsh Jakkula |
| Course Date: 08/09/25 |
|  |

**Table of contents**

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

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

[Day 2: Task 2 3](#_Toc152114794)

[Day 2: Task 3 4](#_Toc257844391)

[Day 3: Task 1 4](#_Toc1014152162)

[Day 3: Task 2 5](#_Toc1498274088)

[Dataset: 5](#_Toc1056274673)

[Step 1: Create a Pivot Table 5](#_Toc782776295)

[Step 2: Use the SWITCH Function 5](#_Toc365195726)

[Submission: 6](#_Toc485671904)

[Day 3: Task 3 6](#_Toc1856180793)

[Day 4: Task 1 7](#_Toc381189142)

[Course Notes 9](#_Toc1368242635)

[Additional Information 10](#_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 | The Data Protection Act 2018 is the current UK law governing how personal data is collected, stored, and processed, ensuring it’s handled fairly, lawfully, and securely to protect individual privacy. It is important to protect people’s privacy and ensure personal data is handled responsibly in the digital age.  For example: A hospital that collects patient’s medical records for treatment purposes should: 1. Obtain consent from the patient 2. Store the data securely 3. Limit access to the data only to authorized staff 4. Allow patients to access and correct their personal data if it isn’t accurate 5. Not to share patient information with third parties unless legally required.  A breach of the data protection act can lead to a fine, reputational damage, and possible legal action with sanctions. |
| GDPR | The General Data Protection Regulation is a European Union regulation that regulates how organisations in the EU/EEA collect, use, store and share personal details, while also protecting individual’s personal data and give them more control over information.  It is necessary so that companies operating in the EU:  1. Ask consumers for consent before sending marketing material. 2. Showing a cookie consent banner before collecting browsing data 3. Allowing users to download or delete their data upon request 4. Ensuring the data is stored in servers secure as per EU/EEA standard  A breach of the GDPR can lead to heavy fines, legal action, reputational damage, and restrictions on data collection/usage. |
| Freedom of Information Act | The Freedom of Information Act 2000 is an act of the Parliament of the UK that creates a public right of access to information held by public/governing authorities. It aims to promote transparency, accountability, and trust between the government and the citizens of the UK.  The act brings responsibility to public authorities to: 1. Regularly publish updates and information regarding their activities.  2. Ensure members of the public can make requests for information and updates regarding the activities of public authorities.  An example of public authorities under the governance of FOI act in the UK include the NHS, police, government departments and state schools. FOI can be restricted on the case of the information being classified, or filing the request being unreasonably difficult.  A breach of the Freedom of Information Act can lead to public distrust, and the public’s reasonable grounds to take legal action. |
| Computer Misuse Act | The Computer Misuse Act 1990 is an Act of the Parliament of the UK to make provision for securing computer material against unauthorized access or modification. It is necessary to protect computer systems from malicious users who try to either gain unauthorised access, including mere physical contact and the scrutiny of data without any interaction with a computer. This is known as ‘hacking’.  This act incentivizes to: 1. Make individuals strong passwords and two-factor authentication to prevent unauthorized access 2. Have organisations monitor systems for hacking attempts and reporting breaches to relevant authorities 3. Prohibiting installation of malware or malicious software on company devices 4. Restricting access to sensitive information to only authorized authorities. 5. Logging network activity to detect and prevent cyber-attacks  A breach of the Computer Misuse Act can lead to fines, criminal charges or imprisonment depending on the severity of the breach. |

# 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 | Dataset Source:  <https://www.kaggle.com/datasets/matthieugimbert/french-bakery-daily-sales>  Since this is too large of a dataset for MS Excel to handle, I will be making use of the data between breakfast hours (9AM to 11AM) for the date 2021-01-12  I used the ‘Find and Replace’ action to replace the comma (,) fields in the price with decimal (.)  I have also formatted the cells according to date, time and currency in Euros, while also making use of Name Manager      Using the table created, the items can easily be arranged according to time, quantity sold, sales revenue.  I highlighted the most and least item quantity sales with green and red. The only way to properly figure out the most and least popular items would be via Pivot Tables. |

# 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? | Germany has customers in the adult demographic, as the data depicts that the only customers in Germany are within the adult range. |
| What country has sales in all markets? | Australia has sales across all markets by age and gender demographic. |
| What are the most profitable markets by country, age group, and gender? | Australia is the most profitable across all categories, while the US is more profitable compared to Australia in the young adults and adult demographic. However, the US is struggling having no customers in the youth market whatsoever. |
| Any other findings? | The UK is the second-most country to have customers across all markets despite being struggling with total sales in contrast to the US. |

# 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? | I would start by preparing a presentation that includes and introduction, body and conclusion, to ensure the delivery is clear and concise.  This would enable me to help visualize the notion I’m presenting based on data. |
| What tools would you use for the delivery? | I would make use of flash cards or short notes of keywords and phrases that I intend on elaborating on. This allows me to recollect the order of points while maintaining full attention towards the board.  The PowerPoint presentation would include data visualizing graphs and charts produced through Excel or alternate visualizing software. |
| What is prospecting and why would you complete this before your delivery? | Prospecting is a process of identifying and understanding the audience before presenting analytical findings.  This is done to put forward concerns of how the loss of a focused demographic of customers may financially impact the company.  Prospecting prior to a well-tailored delivery establishes shared motivations behind the changes necessary. |
| Tell me best practices for public speaking and providing updates to senior leaders | One would have to avoid reading completely off the notes prepared as they intend to be short phrases to recollect.  It’s essential to face and look towards the board members individually at intervals while occasionally referring to the PowerPoint presentation behind me.  It is necessary to make sure that the presentation slides are more visually formal as opposed to raw data which can be overwhelming and convoluted to present.  And above all, an ideal speaker would have good posture and a concise tone for delivery. |
| What will you show the board in your delivery? | The aim of this presentation would be to identify the poor customer retention and understand the cause behind customers leaving the company prior to the renewal.  This could be done via line charts representing the total customers remaining by the end of each quarter in a year, with annual price changes labelled. |
| How will you articulate the changes that are needed? | I would make use of data visualization to present the trends and relation between customer retention and renewal price increases.  The likely reason for the loss of customers would be the variable price changes by the end of the 12-year point.  To rectify this, I would need to propose a fixed-price guarantee after the 12-year term, or alternatively explore loyalty discounts on customers heading towards the 12-year point. |
| Provide a list of online resources and videos that will support your preparation for public speaking | <https://youtu.be/Ns_z4wEtdRM?si=B4I5TXs2RpZysDBy>  <https://www.youtube.com/shorts/tWeSCB88Kow>  <https://www.youtube.com/shorts/jVBTeuzNi2g>  <https://youtu.be/jizZKNnx9wA?si=EViuN8pYWSqzqxSU> |
| Evaluate tools that provide visualisation.  Tell me what they are.  Tell me what you would choose when delivering your presentation and why | MS Excel is a software that would allow me to organize and process data. It provides minimalistic tools for data representation.  On the other hand, Tableau and MS Power Bi offer more vast data visualization techniques (e.g. Geography heat-maps) which assist transform large complex datasets into interactive dashboards.  Ideally I would use MS Excel for quick data organization and Tableau for visualization |

|  |
| --- |
| **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:

|  |
| --- |
| Data’s 3 Vs: Volume, Velocity and Variety.  Quantitative vs Qualitative Data  Nominal vs binomial data  Mean, median, mode, range, variance, standard deviation, inferential  Mean = sum / number of numbers  Median = middle number (if even datasets, then average the two middle numbers)  Mode = value(s) with most occurrence  Range = span of data = largest number – smallest number  Cell reference: =X0  Absolute Reference: =$X$0: shortcut key f4  Reference from another sheet: =SheetName!X0 (do $ for absolute reference) OR alternatively reference using name box: =NameBox  Formulas TAB: Name Manager Select column from header and then Create from Selection  **Select column with ctrl + shift + down arrow key (sometimes PGdown)**  If you want to name a range of values just select the range and rename in Name Box  **Select full range = win + ctrl + A**  Create Table:  To create table: select range and then Insert 🡪 Table  To change table back to normal: Table Design 🡪 Tools 🡪 Convert to Range  Filter Data:  Select the Column then: Data 🡪 Sort and Filter  Multi-Level Sorting: Select range then Data 🡪 Sort 🡪Add Level  Filtering: Select range then Data 🡪 Filter 🡪 Then filter via column header  Data type: Home 🡪 Number 🡪 Change from general to required  Commission = Revenue \* Rate  SUM, AVG, COUNT, MIN, MAX  Greek letter sigma = SUM  To use these functions, select column then Home 🡪 Editing 🡪 Function (result will be at the bottom)  Make sure its : and not ,  To show formula: Formulas 🡪 Formula Auditing 🡪 Show Formulas  Excel Functions  SUM: Calculate the total values  SUMIF(s): Calculate the total sales for a specific category/condition  AVERAGE: Calculate the average from the values  COUNT: Number of cells that have numbers  COUNTA: Number of cells that have numbers and text  COUNTBLANK: Number of cells that are blank  COUNTIF: Number of cells that meet a condition  COUNTIFS: Number of times multiple conditions are met  UNIQUE: List of unique values in a range  IF: brings result based on certain conditions  <https://www.kaggle.com/>  Data visualization ^  SWITCH: Similar to IF function but cannot select range, individual absolute referencing  **VLOOKUP:**  Column index is the column order you want from the selection  Lookup column must be the first column since it cannot look to the left side of the selection range.  **XLOOKUP**  Lookup value is what you’re searching for  The array is the list  Return array is the result  Match mode set to False  **CONCAT(ENATE)**  To combine text from two cells  Data validation: Restrict the user with options via drop-down menu list  Text to column: Splits the column content into two (make sure they’re not in formulas, copypaste as values)  Importing and blending data: Done via power query from Get Data button |
| **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.**