Data Technician

Name: Thandeka Mukazi

Course Date: Monday 12th May – Thursday 15th May 2025

Table of contents 2 Day 1: Task 1 2 Day 2: Task 1 2 Day 2: Task 2 3 Day 2: Task 3 4 Day 3: Task 1 4 Day 3: Task 2 5 Dataset: 5 Step 1: Create a Pivot Table 5 Step 2: Use the SWITCH Function 5 Submission: 6 Day 3: Task 3 6 Day 4: Task 1 7 Course Notes 9 Additional Information 10

Day 1: Task 1

Please research and complete the below boxes on common 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

What is it?

 A law designed to project personal data stored on computers or in an organised paper filing system Gives individuals control over how their personal data is used by organisations, businesses or the government

Why is it important?

- Protects privacy
- Gives control to the individual
- Prevents misuse and harm
- Ensures organisations are held accountable
- Fosters a relationship of trust

Real world example of how it can be followed

- Schools storing personal information about its students
- Only certain faculty members who need to know certain information can access sensitive information

How does it impact working with data?

- Compliance: Organizations must follow laws governing data collection and usage, requiring clear privacy policies and consent.
- Resources: Implementing data privacy measures involves cost for legal compliance and technologies
- Trust: Proper data management builds consumer trust and loyalty.

What if it is breached?

- **Legal Penalties:** Organizations may face significant fines and legal actions.
- **Reputation Damage:** Breaches can harm customer trust and brand reputation.
- Operational Disruptions: Managing breaches diverts resources and can impact business operations.

GDPR

What is it?

GDPR= Law from EU how personal data of individuals is collected, stored and processed and shared.

Why is it important?

- Protects individual privacy
- Fosters transparency and trust
- Accountability of companies
- Promotion ethical data use

Real world example of how it can be followed

- A company that wants to send promotional emails to consumers within the EU
- Emails including unsubscribe links so customers can withdraw consent

How does it impact working with data?

- Organisations must have a legal reason to process personal data such as consent or public task
- Limits data collection (limited to collecting data that is necessary for specific purposes)
- Emphasis on data accuracy and quality
- Requires transparency and accountability
- Gives people data rights
- Must give data breach notifications
- Compulsory security measures must be put in place to protect personal data within organisations

Consequences of GDPR Breach:

- Fines: Up to €20 million or 4% of global annual turnover (whichever is higher), for serious violations.
- Reputational Damage: Loss of trust from users, customers, and partners.
- Legal Action: Data subjects may sue for damages if their rights are violated.
- Operational Disruption: Investigations by Data Protection Authorities (DPAs) may require audits, data access restrictions, or even temporary bans on processing data.
- Mandatory Notifications: You must report serious breaches within 72 hours to a supervisory authority and sometimes to affected individuals.

Freedom of Information Act 2000 (FOIA)

What is it?

- UK legislation that grants the public the right to access information held by public authorities,
- Examples: government departments, local councils, the NHS, schools, and police forces.
- Aims to promote transparency and accountability in public administration

Why is it important?

- Promotes transparency and trust
- Upholds the authority of the public
- Accountability of governments
- Improves public services

Freedom of Information Act

Real world example of how it can be followed?

- Journalist requesting data from a local education authority on school exclusions over the last 5 years with information related to ethnicity, reason for exclusion and year
- If information isn't exempt, it is sent to journalist in a clear format
- Any refusal must be given with reason

How does it impact working with data?

- Influences how public bodies manage and handle data:
- Data Management: Authorities must maintain organized records to comply with potential information requests.
- Transparency: Encourages the sharing of data, fostering trust and accountability.
- Exemptions: Some information may be withheld if its release could harm national security, law enforcement, or personal privacy.
- Data Security: Authorities must protect data from loss or alteration and ensure personal data is handled according to privacy laws.

- Proactive Disclosure: Many public bodies must release certain types of information regularly, reducing the need for individuals to ask for it.

What happens if it is breached?

- Enforcement Actions: The Information Commissioner's Office (ICO) can issue decision notices requiring authorities to comply with the Act .
- Investigations: The ICO can investigate instances of criminal destruction, alteration, or concealment of information.
- Appeals: Individuals can appeal decisions to the Information Rights Tribunal .
- Criminal Penalties: In severe cases, such as intentionally destroying or concealing information to avoid disclosure, individuals can face criminal charges, including fines or imprisonment.
- Reputation Damage: Failing to comply with FOIA can harm the reputation of the public authority, as it may be seen as untrustworthy or secretive.

What is it?

The **Computer Misuse Act 1990** is a law designed to protect computer systems and data. It criminalizes activities such as:

Computer Misuse Act

- Unauthorized access to computer systems (hacking)
- Unauthorized modification of data (e.g., introducing viruses)
- Using computers to commit further offenses (e.g., fraud or identity theft)
- Unauthorized acts that impair the operation of a computer (e.g., denial-of-service attacks)

Why is it important?

- Facilitates punishment of cybercriminals
- Prevents/minimises damage to systems
- Sets legal standard for digital/online behaviour

- Protection against unauthorised access

Real life example of how it would be followed

- A teenager hacking into a school database to find upcoming GCSE papers with the intent to leak/sell them
- Signals unauthorised access and intent to commit further illegal activities
- May result in fines, prison/young offenders etc

How does it impact working with data?

When working with data, especially in professional or academic settings, the Act means:

- You must only access systems and data you are authorized to use
- You should not share login credentials or use someone else's
- You must avoid altering or deleting data unless you have permission
- You need to ensure data is secure, and not vulnerable to unauthorized access

Organizations must also ensure staff are trained and systems are protected from misuse.

What if it is breached?

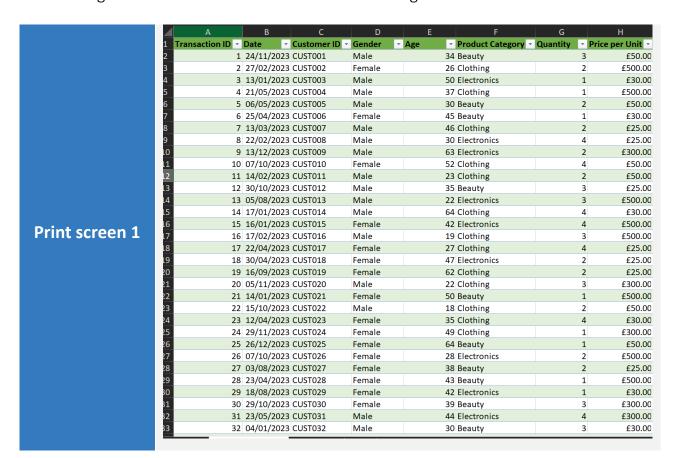
Breaching the Computer Misuse Act is a **criminal offense**. Consequences can include:

- Fines of thousands of pounds
- Prison sentences (up to life imprisonment for serious offenses)
- Loss of job or career prospects
- Damage to your reputation and potential civil lawsuits

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:

- In the sheet 'retail_sales_dataset' add all available data between columns A H into a 'table'
- 2. Using the 'filter' function, filter 'Age' to 'largest to smallest'
- 3. Using the 'SUM' function, show me the commission total in cell 'P10'
- 4. Using the 'AVERAGE' function, show me the average commission in cell 'P11'



	′_Age ↓↓	
	64 (
	64 I	
	64 (
	64 I	
	64 I	
	64 (
	64 I	
	64 (
	64 I	
	64 I	
	64 I	
	64 (
	64 I	
	64 I	
	64 I	
Print screen 2	64 I	
	64 I	
	64 I	
	64 (
	64 I	
	64 (
	64 (
	64 (
	64 I	
	64 (
	64 I	
	64 (
	64 (
	64 (
	64 I	
	64 I	
	63 I	

Print screen 3	Total Com	mission	
		£6,840.00	
Print screen 4	Average Cor	Average Commission	
	£	6.84	

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:

Student name	English	Mathematic:	Science	Average	Highest score	
Carol	75	85	85			
Ted	80	75	90			
Khan	85	75	80			
Harry	80	70	80			
Sarah	80	70	80			
John	65	80	70			
Linda	90	50	70			
Edward	55	80	60			
Mary	55	70	65			
Thomas	55	30	65			
Task						

- 1) Apply filter and sorting to show the best students in each subject.
- 2) Calculate the average for all students and fill into Column E. (Use formula)
- 3) Using the =MAX fucntion, tell me what the students highest score was in column F.
- 4) Apply filter and sorting to show the best student in this classroom by average.
- 5) Apply filter and sorting to show the best student in this classroom by highest score.
- 6) Use conditional formatting to clearly identify the highest and lowest average scores

Print screen 1

1) Linda is the best at English and Mathematics. Ted is the best at Science.

Student name	Englisł↓↓	Mathematic 🕶	Science	Average	Highest scor ▼
Linda	90	85	70	81.666667	90
Khan	85	80	80	81.666667	85
Ted	80	80	90	83.333333	90
Harry	80	75	80	78.333333	80
Sarah	80	75	80	78.333333	80
Carol	75	70	85	76.666667	85
John	65	70	70	68.333333	70
Mary	55	70	65	63.333333	70
Thomas	55	50	65	56.666667	65
Edward	55	30	60	48.333333	60

Student name	Englisł 🔻	Mathematic 🔻	Science↓↓
Ted	80	80	90
Carol	75	70	85
Khan	85	80	80
Harry	80	75	80
Sarah	80	75	80
Linda	90	85	70
John	65	70	70
Mary	55	70	65
Thomas	55	50	65
Edward	55	30	60

Α	В	С	D	Е	F
Student name	Englisł 🔻	Mathematic -	Science	Average↓↓	Highest scor ▼
Ted	80	80	90	83.333333	90
Linda	90	85	70	81.666667	90
Khan	85	80	80	81.666667	85
Harry	80	75	80	78.333333	80
Sarah	80	75	80	78.333333	80
Carol	7 5	70	85	76.666667	85
John	65	70	70	68.333333	70
Mary	55	70	65	63.333333	70
Thomas	55	50	65	56.666667	65
Edward	55	30	60	48.333333	60

2) The averages for students calculated; filled in Column E.

3/4/5) Students highest scores displayed in column F. Ted is the best student in terms of average and highest score.

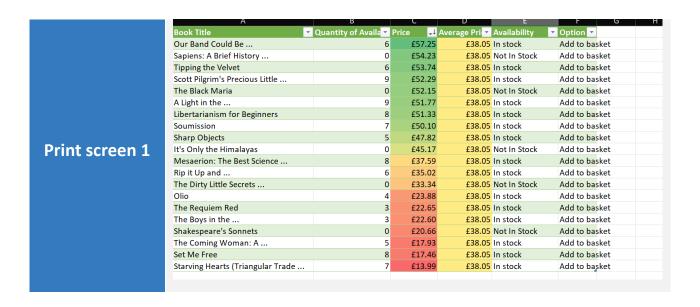
Student name	Englisł 🔻	Mathematic 🔻	Science	Average ▼	Highest scor ↓↓	
Ted	80	80	90	83.333333	90	
Linda	90	85	70	81.666667	90	
Khan	85	80	80	81.666667	85	
Carol	75	70	85	76.666667	85	
Harry	80	75	80	78.333333	80	
Sarah	80	75	80	78.333333	80	
John	65	70	70	68.333333	70	
Mary	55	70	65	63.333333	70	
Thomas	55	50	65	56.666667	65	
Edward	55	30	60	48.333333	60	

Α	В	С	D	Е	F
Student name	Englisł 🔻	Mathematic 🕶	Science	Average ▼	Highest scor ↓↓
Ted	80	80	90	83.333333	90
Linda	90	85	70	81.666667	90
Khan	85	80	80	81.666667	85
Carol	75	70	85	76.666667	85
Harry	80	75	80	78.333333	80
Sarah	80	75	80	78.333333	80
John	65	70	70	68.333333	70
Mary	55	70	65	63.333333	70
Thomas	55	50	65	56.666667	65
Edward	55	30	60	48.333333	60

6) Used conditional formatting to identify the highest and lowest average scores. Green are the highest, amber/yellow are the middle and the red/deep orange are the lower average scores.

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!



Day 3: Task 1

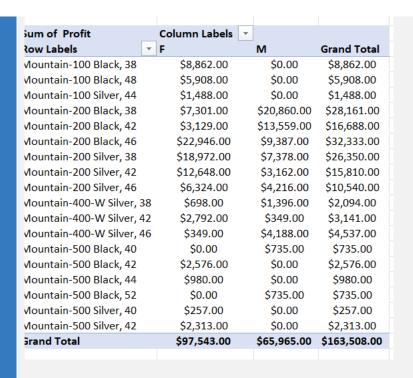
Please download the dataset 'Day_3_Task_1_Bike_Sales_Pivot_Lab.xlsx' from here.

The lab instructions can be found <u>here</u>. 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:



In which markets does Germany have customers?	Germany has customers in the following markets (I wasn't sure how to interpret the question to be honest): - Mountain-200 Black, 38 - Mountain-200 Black, 46 - Mountain-200 Silver, 38 - Mountain-200 Silver, 42 - Mountain-200 Silver, 46
What country has sales in all markets?	None of them have sales in all markets but that could be me totally misinterpreting the question.
What are the most profitable markets by country, age group, and gender?	Country: Australia = Mountain-100 Black, 38 Canada = Mountain-200 Silver, 38 France = Mountain-200 Black, 38 Germany = Mountain-200 Silver, 46 United States = Mountain-200 Black, 46 United Kingdom = Mountain-200 Black, 38 Age Group: Youth = Mountain-200 Silver, 46 Young Adults = Mountain-200 Silver, 38 Adults = Mountain-200 Black, 46 Gender: Males = Mountain-200 Black, 38 Females = Mountain-200 Black, 46



Other Insights?

When it comes to the mountain-100 Black, 38, Mountain-100 Black 48, Mountain-100 Silver 44, Mountain-500 Black 44, Mountain-500 Black 42, Mountain-500 Silver, 40 and Mountain-500 Silver, 42:

- There was no profit gained within the male category for these particular markets
- This points to males being less likely to purchase these particular products
- Marketing efforts could be used to find out why and change their marketing strategy based on that feedback

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:

Co	unty	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:

- o For sales greater than 600: "High"
- o For sales between 300 and 600: "Medium"
- o 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.
 - o Please paste your completed work below

Row Labels	Sum of Sales Volume
□ Cornwall	1100
Laptops	700
Printers	400
□ Durham	550
Laptops	250
Printers	300
□ Essex	1100
Printers	800
Smartphones	300
■ Greater Manchester	1000
Laptops	400
Smartphones	600
□ Lancashire	750
Laptops	600
Smartphones	150
■ Yorkshire	700
■ Yorkshire Laptops	700 500
Laptops	500

Print screen 1

• A completed Pivot Table summarising sales by county and product.

County	Product	Sales Volume	Sales Volume Catgeory
Yorkshire	Laptops	500	Medium
Yorkshire	Smartphones	200	Low
Cornwall	Laptops	700	High
Cornwall	Printers	400	Medium
Lancashire	Smartphones	150	Low
Lancashire	Laptops	600	Medium
Essex	Printers	800	High
Essex	Smartphones	300	Medium
Durham	Laptops	250	Low
Durham	Printers	300	Medium
Greater Manchester	Smartphones	600	Medium
Greater Manchester	Laptops	400	Medium

• A new column in the dataset categorising products by sales volume using the SWITCH function.

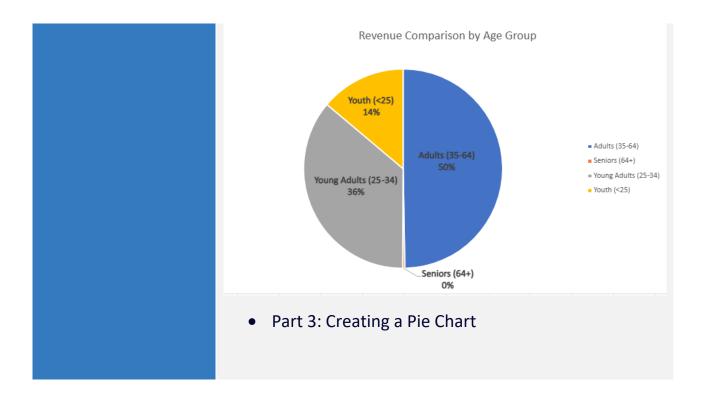
Day 3: Task 3

Please download the dataset 'Day_3_Task_3_Bike_Sales_Visualisations_Lab.xlsx' from here.

The lab instructions can be found <u>here.</u> 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:





Day 4: Task 1

You have been asked to deliver your analysis findings to the board of directors, within 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 go through a planning stage first. This would include:

- To clarify and highlight what the purpose of the presentation is. Maybe in that, enquire about what the board of directors are interested in finding out. In this case, to showcase a customer retention issue at the 12 month mark and what will be presented would align with what the directors are expecting (e.g. just facts or also recommendations?)
- Know the audience well and plan the presentation with that in mind. In this case, the use of nontechnical language may be advisable unless they specifically ask for technical deep dives and insights
- Choosing how to structure the presentation. So if I am using a PowerPoint presentation, I would order

it similar to this: Introduce the problem + why it is important, present the key insights, implications of the problem, recommendations and actionable next steps

Preparing for delivery on the public speaking and communication side:

- Ensure that I know my material inside and out which will help me in my answering of questions

- Practicing out loud, maybe on a practice audience who are similar to the target demographic or can role play as board of directors
- Preparing presentation slides with a balance of information and visuals – keeping it minimalistic and relevant
- Prepare evidence as potential answers to pushback questions

What tools would you use for the delivery?

- Microsoft PowerPoint presentation
- Simple speaker notes that will aid me what to talk about and when
- Excel produced graphs
- A clicker for smooth navigation (would depend if the presentation is in person or remote)

What is prospecting and why would you complete this before your delivery?

Prospecting is the process of identifying and reaching out to potential customers/clients who could take an interest in a product, service or something you are offering. It is key when delivering a presentation that is aimed at influencing decision-makers such as a board of directors.

Why I would complete this before my delivery?

- To enable you to tailor my delivery in a way that matters to the directors, having identified the type of people I will be speaking to
- I can prioritise what you focus on so that it aligns to the directors

Building credibility beforehand Can anticipate objections or questions - Can build focused as well as relevant recommendations/actionable insights Some best practices for public speaking: Knowing your audience Being clear and concise Using visuals intelligently Engaging with eye contact and body language Prepare for questions Tell me best practices for public speaking and providing For providing updates to senior leaders: updates to senior Lead with headlines and key insights leaders Keep a focus on outcome and use data to back up points Keep it brief and structured Stay professional and neutral Leave time for discussion Summary slide – key findings, why it matters and recommended action Line/bar chart showing retention rates over time etc Impact on the business – like revenue loss, What will you show potential gain if retention improves the board in your Using the voice of the customer – outlining quotes delivery? or surveys that highlight dissatisfaction and reasons to leave - Competitor insight Proposed solutions – like improved communication and show the potential impact Next steps (what is needed going forward)—further research, budget etc

How will you articulate the changes that are needed?

- To state and reiterate the problem clearly
- Highlight and explain the business impact
- Note the root causes
- To talk about necessary changes that are relevant
- Overall, link back to the strategic goals of the organisation

Provide a list of online resources and videos that will support your preparation for public speaking

https://www.youtube.com/watch?v=rTRfNQFY4ao https://www.youtube.com/watch?v=i5mYphUoOCs https://www.youtube.com/watch?v=bsxJVgb6Kls

https://medium.com/learning-data/5-public-speaking-tips-for-data-analysts-how-to-calm-your-nerves-and-crush-that-presentation-3cf69c0e062e

Evaluate tools that provide visualisation.

Microsoft Excel

- Pro = Ideal for quick charts such as bar, line and pie charts
- Con = Limited in design customisation

Tell me what they are.

Power BI

- Pro = Real time updates and great integration of excel, databases, SharePoint etc
- Con = Power BI pro required to access full sharing functions

Tell me what you would choose when delivering your presentation and why

I would use Microsoft Excel as a visualisation tool because I can prepare visuals quickly and excel tends to be a format that most people are familiar with.

- I could use line/bar charts to show retention over
 12 months
- Pivot tables/charts to easily breakdown leaving (churn) rates by different categories such as customer type

- I could display these within a PowerPoint presentation
- Maybe a summary predictive/what if analysis chart or table to showcase impact based on recommendations/actionable insights being carried out

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:			

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.