DATA ANALYSIS – PORTFOLIO 2

With a diverse background spanning from borehole operations technician to elementary classroom teacher, cashier, front-desk banking officer, and now a data scientist, I have embarked on a journey to continually prove my worth and quantify my relevance in various fields.

My tenure as an intern at Lufemos Consult Nig Ltd has equipped me with valuable skills in data analytics using Microsoft Excel spreadsheets, SQL Server Management Studio for database management, Power BI for visualization, and Python programming along with data science libraries like NumPy, Pandas, and Matplotlib.

The acquisition of this set of skills for data science tasks and a future of endless learning has driven a sense of career fulfilment within me and I am passionate and optimistic of a long and successful career in data science.

PROFESSIONAL BACKGROUND

Armed with a bachelor's degree in Applied Geology, I possess practical knowledge in water borehole construction techniques and have experience as a sales representative, providing me with a well-rounded understanding of diverse operational environments, including the banking sector. Motivated by a desire for a sustainable and dynamic career that aligns with my objectives, I transitioned to data science, accumulating two years of experience in utilizing tools such as Microsoft Excel, SQL Server Management Studio (SSMS) Workstation, Power BI, and Python for data analysis, visualization, and reporting.

My proficiency in Python programming language for data science tasks with libraries such as Pandas, and NumPy, drives me to contribute to clearer data-driven decision-making processes through insightful reports. With dedication and consistency, I am committed to leveraging my skills and experiences to make meaningful contributions in the field of data science and beyond**.**

TABLE OF CONTENT

INTRODUCTION

PROJECT BACKGROUND

This project is an end-of-course project, at the end of this course, I've gained significant expertise in Microsoft Excel, mastering techniques to organize data for analysis effectively. This project entails crafting a comprehensive summary of sales figures for Adventure Works, comparing first-quarter results between 2022 and 2023. The objective is to demonstrate proficiency in creating formulas to calculate total sales, percentage increases, and month-by-month breakdowns. By leveraging Excel's functionalities, I aim to deliver a detailed report that provides valuable insights into sales performance, aiding in strategic decision-making and cleaning data for further analysis. This endeavour showcases my ability to manipulate data and generate meaningful insights, highlighting my competence in data analysis and reporting.

PROBLEM

The task involves restructuring and analyzing sales data for Adventure Works, focusing on the first quarter's performance for two consecutive years. The existing dataset contains information on product sales in a raw and disorganized form, including wholesale and retail prices, quantities sold, and order dates. The objective is to create comprehensive cleaned-up data that highlights total sales, monthly breakdowns, and year-over-year comparisons and is ready for further analysis as well.

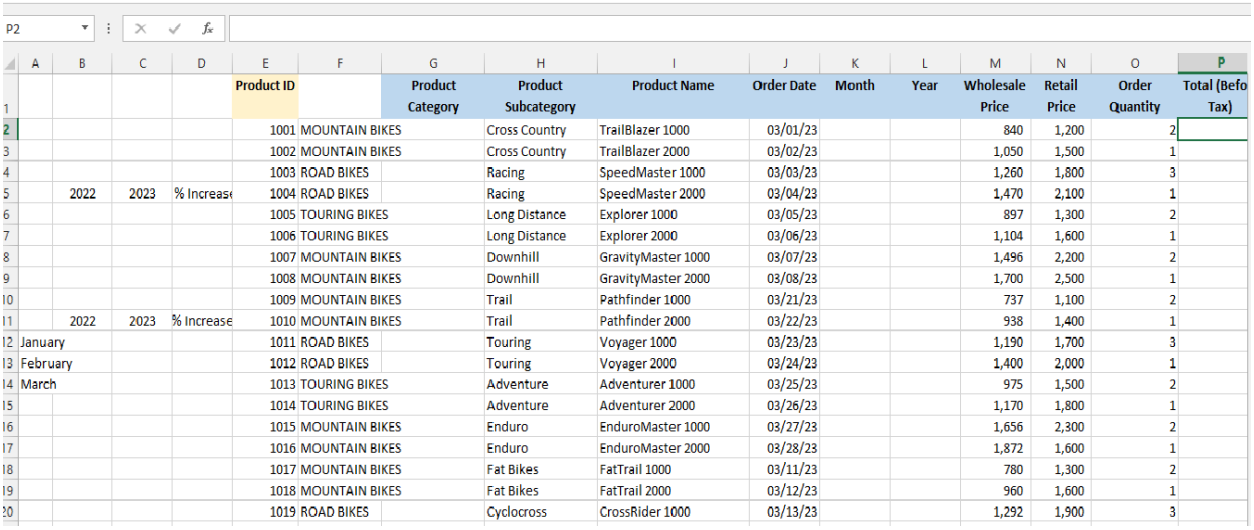
DESIGN

The process of achieving a perfect summary for the quarter one report involves narrowing down blank spaces within the columns of the sheet by the use of the TRIM function in Excel, cleaning up improper headings, adjusting fonts where necessary, background colours and also merging and centring text where necessary.

The use "PROPER" function for accurately organizing text made the worksheet look more presentable. Other clean-up functions used for this exercise included the use of SUMIF and percentage functions to add more meaning to the end product of the data-cleaning exercise.

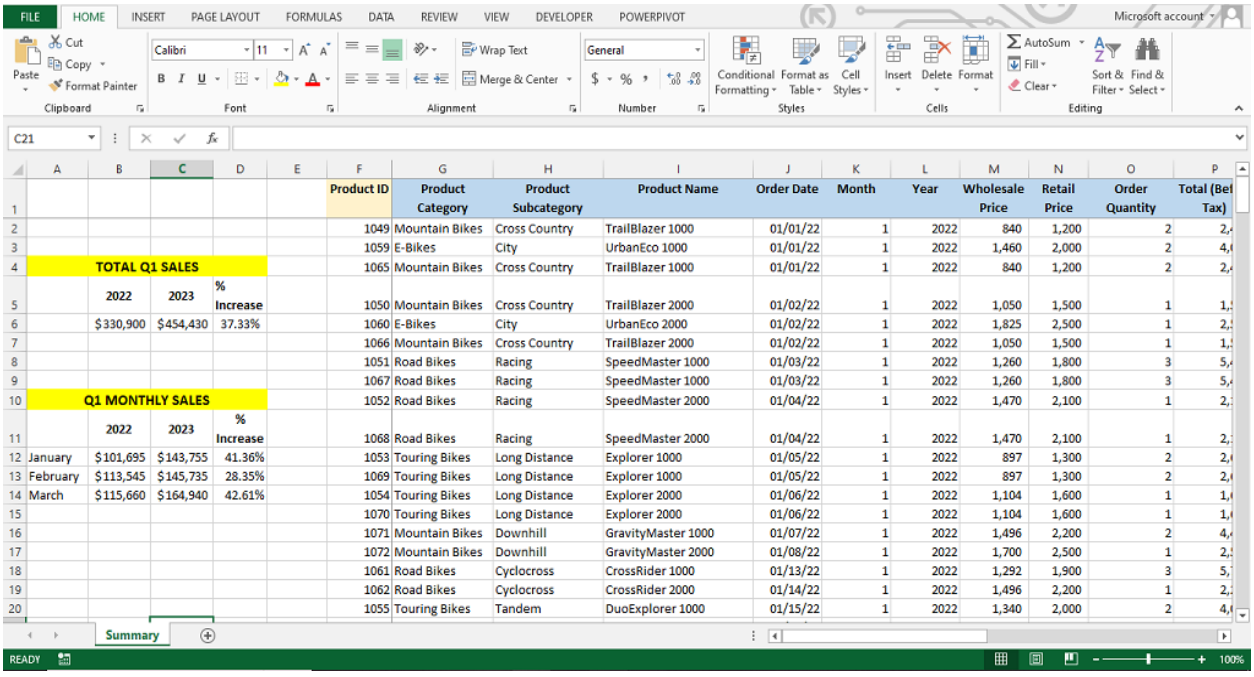
At the end of the day, the sales data for Adventure Works was cleaned and ready for further analysis/presentation.

VISUALS OF EXCEL WORKSHEET BEFORE AND AFTER CLEANING



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1					Product ID		Product Category	Product Subcategory	Product Name	Order Date	Month	Year	Wholesale Price	Retail Price	Order Quantity	Total (Befo Tax)
2					1001	MOUNTAIN BIKES		Cross Country	TrailBlazer 1000	03/01/23			840	1,200	2	
3					1002	MOUNTAIN BIKES		Cross Country	TrailBlazer 2000	03/02/23			1,050	1,500	1	
4					1003	ROAD BIKES		Racing	SpeedMaster 1000	03/03/23			1,260	1,800	3	
5		2022	2023	% Increase	1004	ROAD BIKES		Racing	SpeedMaster 2000	03/04/23			1,470	2,100	1	
6					1005	TOURING BIKES		Long Distance	Explorer 1000	03/05/23			897	1,300	2	
7					1006	TOURING BIKES		Long Distance	Explorer 2000	03/06/23			1,104	1,600	1	
8					1007	MOUNTAIN BIKES		Downhill	GravityMaster 1000	03/07/23			1,496	2,200	2	
9					1008	MOUNTAIN BIKES		Downhill	GravityMaster 2000	03/08/23			1,700	2,500	1	
10					1009	MOUNTAIN BIKES		Trail	Pathfinder 1000	03/21/23			737	1,100	2	
11		2022	2023	% Increase	1010	MOUNTAIN BIKES		Trail	Pathfinder 2000	03/22/23			938	1,400	1	
12	January				1011	ROAD BIKES		Touring	Voyager 1000	03/23/23			1,190	1,700	3	
13	February				1012	ROAD BIKES		Touring	Voyager 2000	03/24/23			1,400	2,000	1	
14	March				1013	TOURING BIKES		Adventure	Adventurer 1000	03/25/23			975	1,500	2	
15					1014	TOURING BIKES		Adventure	Adventurer 2000	03/26/23			1,170	1,800	1	
16					1015	MOUNTAIN BIKES		Enduro	EnduroMaster 1000	03/27/23			1,656	2,300	2	
17					1016	MOUNTAIN BIKES		Enduro	EnduroMaster 2000	03/28/23			1,872	1,600	1	
18					1017	MOUNTAIN BIKES		Fat Bikes	FatTrail 1000	03/11/23			780	1,300	2	
19					1018	MOUNTAIN BIKES		Fat Bikes	FatTrail 2000	03/12/23			960	1,600	1	
20					1019	ROAD BIKES		Cyclocross	CrossRider 1000	03/13/23			1,292	1,900	3	

Fig 1: Raw data before cleaning



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1						Product ID	Product Category	Product Subcategory	Product Name	Order Date	Month	Year	Wholesale Price	Retail Price	Order Quantity	Total (Bef Tax)
2						1049	Mountain Bikes	Cross Country	TrailBlazer 1000	01/01/22	1	2022	840	1,200	2	2,4
3						1059	E-Bikes	City	UrbanEco 1000	01/01/22	1	2022	1,460	2,000	2	4,(
4		TOTAL Q1 SALES				1065	Mountain Bikes	Cross Country	TrailBlazer 1000	01/01/22	1	2022	840	1,200	2	2,4
5		2022	2023	% Increase		1050	Mountain Bikes	Cross Country	TrailBlazer 2000	01/02/22	1	2022	1,050	1,500	1	1,5
6		$330,900	$454,430	37.33%		1060	E-Bikes	City	UrbanEco 2000	01/02/22	1	2022	1,825	2,500	1	2,2
7						1066	Mountain Bikes	Cross Country	TrailBlazer 2000	01/02/22	1	2022	1,050	1,500	1	1,5
8						1051	Road Bikes	Racing	SpeedMaster 1000	01/03/22	1	2022	1,260	1,800	3	5,4
9						1067	Road Bikes	Racing	SpeedMaster 1000	01/03/22	1	2022	1,260	1,800	3	5,4
10		Q1 MONTHLY SALES				1052	Road Bikes	Racing	SpeedMaster 2000	01/04/22	1	2022	1,470	2,100	1	2,2
11		2022	2023	% Increase		1068	Road Bikes	Racing	SpeedMaster 2000	01/04/22	1	2022	1,470	2,100	1	2,2
12	January	$101,695	$143,755	41.36%		1053	Touring Bikes	Long Distance	Explorer 1000	01/05/22	1	2022	897	1,300	2	2,(
13	February	$113,545	$145,735	28.35%		1069	Touring Bikes	Long Distance	Explorer 1000	01/05/22	1	2022	897	1,300	2	2,(
14	March	$115,660	$164,940	42.61%		1054	Touring Bikes	Long Distance	Explorer 2000	01/06/22	1	2022	1,104	1,600	1	1,(
15						1070	Touring Bikes	Long Distance	Explorer 2000	01/06/22	1	2022	1,104	1,600	1	1,(
16						1071	Mountain Bikes	Downhill	GravityMaster 1000	01/07/22	1	2022	1,496	2,200	2	4,(
17						1072	Mountain Bikes	Downhill	GravityMaster 2000	01/08/22	1	2022	1,700	2,500	1	2,5
18						1061	Road Bikes	Cyclocross	CrossRider 1000	01/13/22	1	2022	1,292	1,900	3	5,7
19						1062	Road Bikes	Cyclocross	CrossRider 2000	01/14/22	1	2022	1,496	2,200	1	2,2
20						1055	Touring Bikes	Tandem	DuoExplorer 1000	01/15/22	1	2022	1,340	2,000	2	4,(

Fig 2: Data after cleaning

CONCLUSION

In conclusion, the aim of having a well-organized worksheet for presentation and further analysis was achieved by the use of logical and percentage functions in Microsoft Excel. This is a very important skill in the field of data analysis as it helps improve the optimization and workability of any data set prior to further processing.

REFERENCE

Preparing data for analysis with Microsoft Excel – by Microsoft.