Masar Excel Bootcamp Final Project.

PRESENT TO:

Professional Association for Statisticians and Data scientist

PRESENT BY:

Rawan Hamdi Aljohani

rawan.h.aljo@gmail.com

CONTENTS

INTRODUCTION:

Microsoft Excel is a powerful tool widely used in business and management. It helps people analyze data, create financial models, track budgets, and visualize trends. Being good at Excel improves decision-making, creates data-driven strategies, and boosts organizational efficiency. Understanding the importance of these skills, the Masar Excel Initiative, led by the Professional Association for Statisticians and Data Scientists, aims to connect academic knowledge with real-world job needs.

The Masar Excel Bootcamp gathers a diverse group of participants to learn advanced Excel skills through hands-on projects and group discussions over 4 weeks. Participants apply their new Excel knowledge to real-world business problems, becoming skilled data analysis experts. This program gives us a chance to show our Excel skills and understanding.

This report is my final project for the Masar Excel Bootcamp. It demonstrates the practical use of Excel tools and techniques learned during the program. Each section highlights how Excel can solve complex business problems through data manipulation, visualization, complex analysis, and automation using VBA.

During the boot camp, I have improved my skills in data analysis and financial modeling. Working on real projects with complex formulas and dynamic data visualizations has made me much better at using Excel. Now, I can solve business problems more efficiently and with better analysis.

I would like to express my gratitude to my supervisor, Omar Alghreeb, for his invaluable help and support throughout this program. His guidance and motivation have played a pivotal role in my development and accomplishments.

ANSWERS

Q1: Add a column named “Profit” as “Cost – Sale", after the column “Sale”. How many days between the first order and the previous order?

In question one, I created an ID starting from A2 to uniquely identify each row in the table. I added a "Profit" column after the "Sale" column. The profit for each row was calculated by subtracting the "Cost" from the "Sale". In the "Q1" sheet, I use conditional formatting to select costs and sales above the average with red color.

	A	B	C	D	E	F	G	H	I
1	ID	date	المدينة	ممثل المبيعات	المنفذ	المنتج	Cost	Sale	profit
2	1	10/9/2005	الجبيل	يحيى	العثيم	ألعاب	373.665	498.22	124.555
3	3	2/18/2006	الرياض	مجيد	لولو هايبر	منزلية	3718.965	4958.62	1239.655
4	4	1/1/2006	الجبيل	يحيى	بندة	منزلية	2791.4493	4731.27	1939.8207
5	5	1/29/2005	جدة	مجيد	بندة	منزلية	185.3816	272.62	87.2384
6	6	1/12/2005	الدمام	عبدالله	العثيم	غذائية	3305.9025	7346.45	4040.5475
7	7	1/19/2006	الدمام	يحيى	لولو هايبر	مكتبية	5358.7672	7880.54	2521.7728
8	8	10/12/2004	حائل	يحيى	العثيم	ألعاب	7133.745	9511.66	2377.915
9	9	3/15/2005	الدمام	مليحة	كارفور	غذائية	3491.8709	8120.63	4628.7591
10	10	2/17/2006	جدة	يحيى	لولو هايبر	مكتبية	1350.063	3000.14	1650.077
11	11	12/14/2005	الجبيل	يحيى	كارفور	غذائية	947.6383	2203.81	1256.1717
12	12	10/15/2004	جدة	غالية	بندة	منزلية	2648.9625	3531.95	882.9875
13	13	12/26/2005	جدة	مليحة	كارفور	منزلية	3309.46	3760.75	451.29

Then I created a table to calculate the average cost and sale for items with costs above the average, providing insights into the performance of higher-cost items.

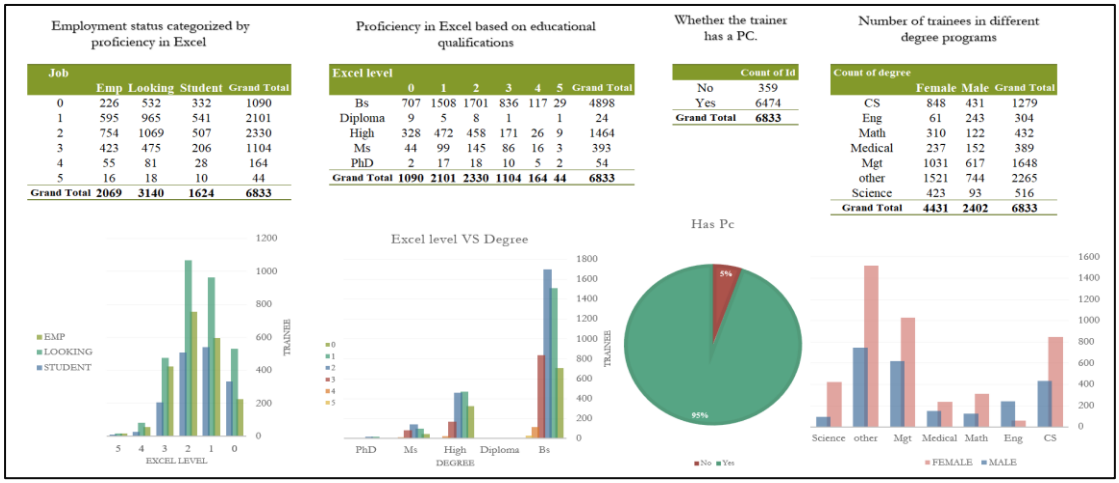
	A	B	C	D	E
1	Avg Cost	Avg Sale	Avg cost only for the cost above the avg	Avg Sale only for the cost above the sale	Days between the first and last order
2	3149.89	5042.72	4943.71	7581.56	4211

Q2: With a proper function, change “Female” to “F” and “Male” to “M” for the entire table in the Q2 sheet. Create at least three pivot tables and provide comments on your findings.

In the "Q2" sheet, I used the IF function to replace "Female" with "F" and "Male" with "M" throughout the table. Then, I created three pivot tables.

1. The first pivot table for Employment status categorized by proficiency in Excel shows that most trainees are looking for job opportunities and have a proficiency level of 2 in Excel. The graph shows that there are fewer trainees with proficiency levels 4 and 5 in Excel compared to those at lower proficiency levels.
2. The second pivot table displays Proficiency in Excel based on educational qualifications and highlights that the majority of trainees held BS degrees we can see from the chart their proficiency level in Excel is between 1 and 2
3. I was confused about their majors. Therefore, I created another pivot table showing the number of trainees in different degree programs to present that Management major students were the top trainees, followed by CS students. Based on this analysis, I recommend that the decision-maker focus on data management, Basic formulas and functions, creating charts and reports, Pivot tables, and Advanced data analysis with pivot tables and formulas.

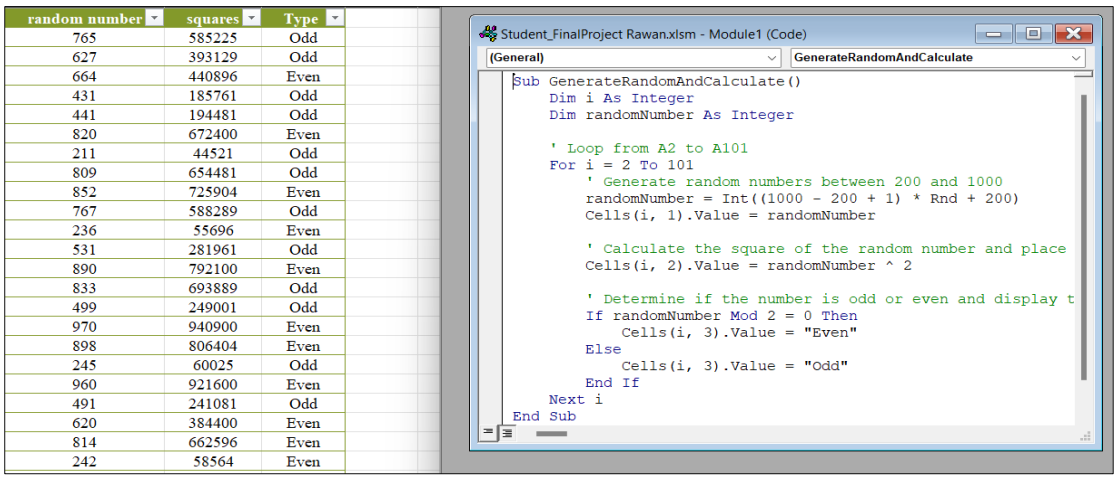
4. The data from the fourth chart shows that 359 trainees did not have access to personal computers. Although this number is relatively small, we can recommend trainees to borrow a PC from a friend or family member.



Q3: Write a VBA code to generate random numbers between 200 and 1000, determine whether each number in column A is odd or even, and display the result in column C.

The VBA code is designed to automate tasks in an Excel spreadsheet. It performs the following actions:

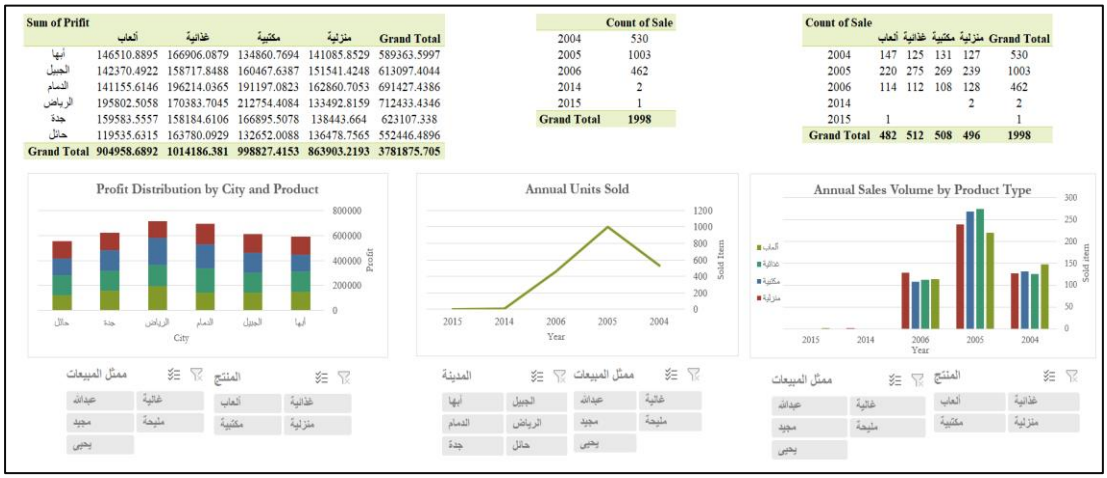
1. **Generating Random Numbers**: Random numbers between 200 and 1000 are generated in column A, from cell A2 to A101. This is achieved using the Rnd function to generate random numbers within the specified range and the Randomize statement to initialize the random number generator.
2. **Calculating Squares**: The squares of the random numbers generated in column A are calculated and placed in column B. This is done by using the exponentiation operator ^ to raise each number to the power of 2.
3. **Determining Odd or Even**: For each number in column A, the code determines whether it is odd or even and displays the result in column C. This is achieved by using the modulo operator (Mod) to check if the number is divisible by 2. If the remainder is 0, the number is even; otherwise, it is od
4. **Execution**: The code is executed by running the "GenerateRandomNumbers" macro from the Excel interface, which automates the process of generating random numbers, calculating their squares, and determining their parity.



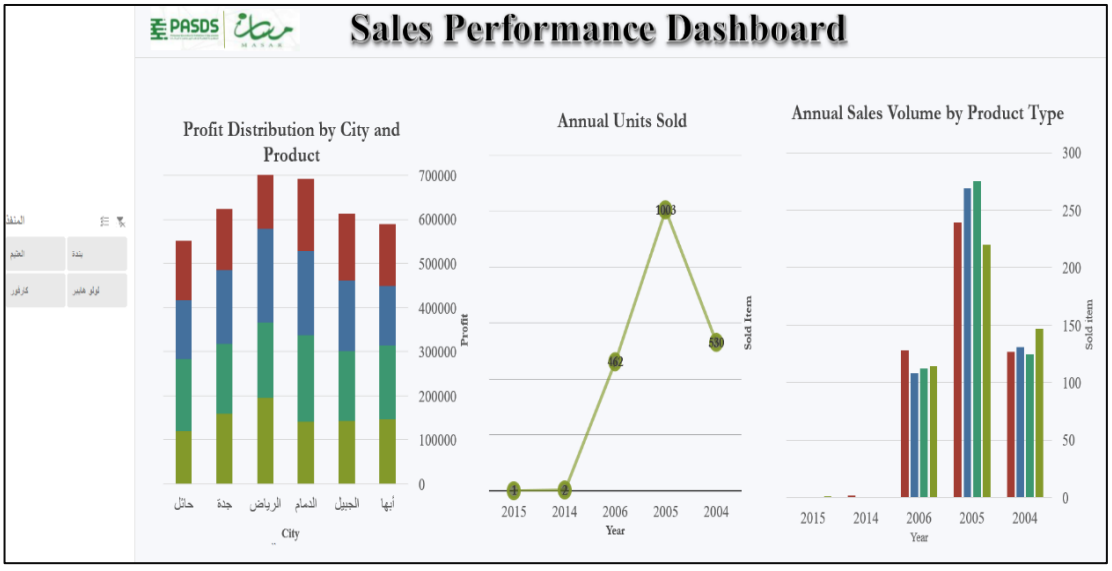
Q4: Create the following pivot tables: a. Sum of profit by City and Product, b. Number of units sold by year. Add two slicers and create a dashboard.

I have created three pivot tables using the sales dataset to analyze different cities in KSA.

1. The first pivot table displays the profit distribution by city and product. I also included two slicers for sales representatives and products, which will help decision-makers assess the impact of all factors on profit.
2. The second pivot table displays the number of units sold annually over 11 years, with two slicers for cities and sales representatives allowing us to analyze their impact on units sold. This chart provides information for making decisions such as marketing strategies, and opening or closing shops.
3. The last pivot table I created was to analyze the number of units sold by product type. I want to find out the best-selling type of product, and for this, I included two slicers for cities and sales representatives. This chart provides more details.



4. Then, I designed a sales performance dashboard featuring three previous graphs and a slicer that allows decision-makers to view all supermarkets' individualized data for them.



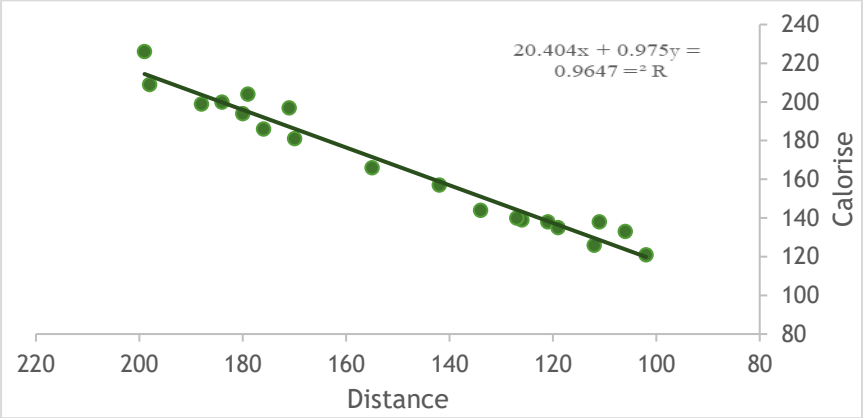
Q5: State the hypothesis for the correlation between distance traveled and calories, and create a scatter plot illustrating the relationship between the two variables.

Hypothesis: we are examining the link between Distance Traveled and Calories Burned. There is a positive relationship, meaning the rate of calories burned increases as the distance traveled increases.

Analysis of the correlation between Distance Traveled and Calories Burned:

CORRELATION COEFFICIENT (R): This measures the strength of the relationship between variables. The relationship between distance and calories is 0.98, indicating a strong positive correlation between variables. in other words, the way we would interpret it is that people who walk long distances tend to burn more calories.

P-VALUE: This points to a significant relationship between variables. and in this case, the P-value is extremely small and indicates a statistically significant relationship.



WRITTEN RESULT:

There is a significant positive relationship between Distance Traveled and Calories Burned, r(11) = 0.98, p < 0.5 (alpha = 0.5) and this strong positive correlation confirms the previous hypothesis.

CONCLUSION

The Masar Excel Bootcamp has been a transformative experience. I've learned many aspects of Excel, from basic formatting and data analysis to more advanced techniques like VBA and pivot tables.

Key Takeaways and Skills Acquired:

- Data Analysis and Visualization: I learned to use Excel for detailed data analysis through practical exercises.
- Automation and Efficiency: This skill is useful for handling large datasets and routine data processing.
- Practical Application through Real-world Projects.

Project Achievements:

- Creation of a Comprehensive Sales Dashboard: I demonstrated my ability to analyze sales data from various angles and present insights through dynamic dashboards.
- Correlational Analysis: I analyzed the relationship between Distance Traveled and Calories Burned, showing my understanding of statistical concepts and their application in Excel.

As I complete this program, I look forward to using these skills in my career, contributing to data-driven decision-making and innovation. I send my best wishes to the founders and supporters of the Masar Excel Initiative, with hopes for the continued success and expansion of this valuable program. I am optimistic about securing my dream job, armed with the knowledge and experience gained from this boot camp.