

ASSIGNMENT 5

1. Student data

Task:

- Create a scatter plot to visualize the relationship between study hours and test scores.
- Generate a histogram to show the distribution of test scores among students.
- Create a bar chart to compare the average test scores of male and female students.
- Visualize the distribution of sleep hours among students using a box plot.
- Create a line chart to track the changes in study hours over time.
- Create a pivot table to summarize the average test scores based on gender and age group.
- Use a pivot table to calculate the total study hours for each gender.
- Generate a pivot table to analyze the average test scores based on study hours and sleep hours.
- Create a pivot table to identify the top-performing students based on test scores.
- Use a pivot table to calculate the median test score for each age group.
- Use a doughnut chart (*With their respective percentages as labels*) to show the distribution of gender of students

Student ID	Name	Age	Gender	Study Hours	Sleep Hours	Test Score
101	Chinedu Okoro	18	Male	5	7	85
102	Ngozi Ibrahim	19	Female	6	8	90
103	Ifeanyi Adeleke	20	Male	4	6	75

Student ID	Name	Age	Gender	Study Hours	Sleep Hours	Test Score
104	Chioma Okafor	21	Female	7	7	92
105	Tunde Bello	19	Male	5	8	88
106	Folake Adebayo	20	Female	6	6	80
107	Segun Ogunlade	18	Male	7	7	87
108	Amaka Udo	19	Female	4	8	82
109	Nnamdi Eze	20	Male	6	7	89
110	Chinwe Ogbonna	21	Female	5	6	78
111	Chukwudi Nwosu	19	Male	6	7	86
112	Adaku Onwuka	20	Female	7	8	94
113	Olumide Ajayi	21	Male	5	7	83
114	Funke Adekunle	19	Female	6	8	88
115	Chika Okafor	20	Female	4	6	72
116	Eze Okoro	21	Male	7	8	91
117	Bukola Alabi	18	Female	6	7	85
118	Kehinde Ajala	19	Male	5	8	84
119	Aisha Yusuf	20	Male	6	7	89
120	Christabel Okoro	21	Female	7	6	90

2. Sales Trends Visualization

Data: Sales

Task:

- Calculate the total revenue generated from sales in each region.
- Calculate a new 'Profit' column for each transaction.
- Create a pivot table to summarize total sales revenue by month and region.
- Use a pivot table to calculate the average price and quantity sold for each product category.
- Create a pivot table to summarize total sales revenue and profit by product category.
- Identify the top-selling products based on the total quantity sold.
- Identify the top-selling products based on the profit.
- Analyze the distribution of sales over different months and quarters.
- Create a bar chart to visualize the profit for each product category.
- Generate a Combo chart to display the trend of total sales over time (*using the new calculated 'month' column*).
- Generate a line chart to display the trend of total sales over time (*using the new calculated 'weekday' column*).
- Construct a pie chart to illustrate the distribution of sales across different regions.
- Create a stacked column chart to compare the contribution (display percentages on the bars) of each product to total sales revenue.
- Generate a bubble plot to visualize the relationship between price and quantity sold for all products.
- Apply data labels and axis titles to enhance readability and interpretation of all the charts.
- Add trendlines to highlight the overall sales trend for each product category.

Date	Product	Category	Price (USD)	Quantity	Customer Name	Region	Cost (USD)	Revenue (USD)
2023-01-05	Laptop	Electronics	800	2	John Smith	North	600	1600
2023-01-10	Smartphone	Electronics	500	3	Emily Johnson	West	300	1500
2023-01-15	Tablet	Electronics	300	5	David Brown	East	200	1500
2023-01-20	Headphones	Electronics	100	10	Sarah Lee	South	50	1000
2023-01-25	Smartwatch	Electronics	200	8	Michael Clark	Midwest	150	1600
2023-02-01	Television	Electronics	1000	1	Lisa Davis	Northeast	800	1000
2023-02-05	Refrigerator	Appliances	1200	2	James Wilson	Southeast	900	2400
2023-02-10	Washing Machine	Appliances	800	3	Emma Martinez	Northwest	600	2400
2023-02-15	Microwave Oven	Appliances	200	4	Daniel Taylor	Central	150	800
2023-02-20	Blender	Appliances	50	6	Olivia Garcia	Southwest	30	300
2023-02-25	Toaster	Appliances	30	8	William Moore	Southeast	20	240
2023-03-01	Coffee Maker	Appliances	100	3	Sophia Nguyen	Northwest	70	300

Date	Product	Category	Price (USD)	Quantity	Customer Name	Region	Cost (USD)	Revenue (USD)
2023-03-05	Vacuum Cleaner	Appliances	150	2	Ethan Anderson	Midwest	100	300
2023-03-10	Sofa	Furniture	500	1	Ava Wilson	Northeast	350	500
2023-03-15	Dining Table	Furniture	300	2	Mia Thompson	Southwest	200	600
2023-03-20	Bed	Furniture	700	1	Noah White	Central	500	700
2023-03-25	Chair	Furniture	150	4	Charlotte Baker	Midwest	100	600
2023-04-01	Desk	Furniture	200	2	Aiden Harris	Southeast	150	400
2023-04-05	Bookshelf	Furniture	100	3	Harper King	Northwest	80	300
2023-04-10	Lamp	Home Decor	50	5	Jackson Lopez	Southwest	30	250
2023-04-15	Rug	Home Decor	80	4	Lily Brown	Central	60	320
2023-04-20	Clock	Home Decor	40	6	Noah Thompson	Northeast	30	240
2023-04-25	Wall Art	Home Decor	70	3	Olivia Taylor	Southeast	50	210

Date	Product	Category	Price (USD)	Quantity	Customer Name	Region	Cost (USD)	Revenue (USD)
2023-05-01	Vase	Home Decor	25	8	Ethan Wilson	Midwest	120	200
2023-05-05	Candle Holder	Home Decor	15	10	Mia Garcia	Northwest	80	150
2023-05-10	Plant Pot	Home Decor	20	7	Ava Thompson	Southwest	90	140
2023-05-15	Photo Frame	Home Decor	10	12	Harper Davis	Central	60	120
2023-05-20	Mirror	Home Decor	60	3	Jackson Moore	Northeast	120	180
2023-05-25	Throw Pillow	Home Decor	30	5	Lily Martinez	Southeast	80	150
2023-06-01	Curtain	Home Decor	40	4	Noah King	Midwest	100	160
2023-06-05	Table Runner	Home Decor	15	8	Olivia White	Northwest	70	120
2023-06-10	Wall Clock	Home Decor	25	6	Ethan Brown	Southwest	90	150
2023-06-15	Sculpture	Home Decor	50	2	Mia Davis	Central	40	100
2023-06-20	Candle	Home Decor	5	15	Ava Thompson	Northeast	50	75
2023-06-25	Wall Shelf	Home Decor	30	3	Harper Lee	Southeast	40	90

3. Data: Employee performance

Task:

- Create a pivot table to summarize the total sales revenue by department.
- Use a pivot table to calculate the average customer satisfaction rating by department.
- Generate a pivot table to analyze the average salary by tenure range (*e.g., 0-2 years, 3-5 years, 6+ years*).
- Design a pivot table to summarize the total actual sales by employee and department.
- Use a pivot table to analyze the maximum performance score by department and tenure range.
- Create a scatter plot to visualize the relationship between employee tenure and actual sales. What relationship is there?
- Generate a bar chart showing the distribution of performance scores among different departments.
- Create a line chart to compare the sales target and actual sales for each employee.
- Generate a pie chart to represent the proportion of employees in each department.
- Plot a Radar chart of the various departments and their total performance scores. (*you will need a pivot table*)

Employee ID	Employee Name	Department	Tenure (years)	Salary (USD)	Sales Target (USD)	Actual Sales (USD)	Customer Satisfaction Rating (%)	Performance Score (out of 10)
101	John Smith	Sales	3	60000	100000	120000	90	8
102	Emily Johnson	Marketing	5	70000	80000	85000	85	7
103	David Brown	HR	2	50000	-	-	-	-
104	Sarah Lee	Finance	4	65000	90000	95000	92	9
105	Michael Clark	Sales	6	75000	120000	115000	88	8
106	Lisa Davis	Marketing	3	60000	750000	70000	80	7
107	James Wilson	HR	4	55000	-	-	-	-
108	Emma Martinez	Finance	7	80000	100000	105000	94	9
109	Daniel Taylor	Sales	5	70000	110000	100000	85	7
110	Olivia Garcia	Marketing	2	50000	70000	60000	75	6
111	William Moore	HR	3	60000	-	-	-	-
112	Sophia Nguyen	Finance	6	75000	95000	90000	90	8

Employee ID	Employee Name	Department	Tenure (years)	Salary (USD)	Sales Target (USD)	Actual Sales (USD)	Customer Satisfaction Rating (%)	Performance Score (out of 10)
113	Ethan Anderson	Sales	8	8500	13000	12500	87	8
114	Ava Wilson	Marketing	4	6500	8000	75000	82	7
115	Mia Thompson	HR	5	7000	-	-	-	-
116	Noah White	Finance	3	6000	8500	8000	88	8
117	Charlotte Baker	Sales	7	8000	14000	13000	90	9
118	Aiden Harris	Marketing	2	5000	7000	6500	78	6
119	Harper King	HR	4	5500	-	-	-	-
120	Jackson Lopez	Finance	5	7000	9000	8500	92	9
121	Emma Smith	Sales	6	7500	12000	11500	89	8

4. Data: Agriculture

Task:

- Use a pivot table to calculate the average yield for each crop type.
- Summarize the total yield generated by each crop type using a pivot table.
- Analyze the average pesticide and fertilizer usage for each category of crops using pivot tables.
- Create a pivot table to examine the correlation between crop yield and soil pH levels. Create a regression model, determine the independent variable and the dependent variable, and make a prediction using the model.
- Generate a pivot table to compare the total revenue generated from different suppliers.
- Create a Combo chart to visualize the relationship between crop yield and temperature.
- Generate a scatter plot to analyze the correlation between rainfall and pesticide usage. Interpret the scatter plot.
- Design a histogram to visualize the distribution of soil pH levels across different farms.
- Develop a box plot to compare the yield distributions of different crop types.
- Use a Tree map to show the average temperature of each crop type.
- Construct a stacked bar chart to visualize the contribution of fertilizer and pesticide usage to total agricultural yield.

Farm ID	Crop	Crop Type	Yield (kg/acre)	Temperature (°C)	Rainfall (mm)	Soil pH	Pesticide Used (kg)	Fertilizer Used (kg)
101	Wheat	Grain	1500	20	50	6.5	20	100
102	Rice	Grain	2000	25	100	6.0	25	120
103	Maize	Grain	1800	22	80	6.2	18	110

Farm ID	Crop	Crop Type	Yield (kg/acre)	Temperature (°C)	Rainfall (mm)	Soil pH	Pesticide Used (kg)	Fertilizer Used (kg)
104	Cotton	Fiber	800	30	120	7.0	30	150
105	Soybean	Legume	1000	28	90	6.8	22	130
106	Barley	Grain	1400	18	40	6.3	15	90
107	Sugarcane	Cane	2500	35	150	6.6	40	200
108	Potato	Root	1200	15	60	5.8	10	80
109	Tomato	Fruit	1800	25	70	6.5	20	120
110	Sunflower	Oilseed	900	28	80	6.7	25	140
111	Peanuts	Legume	1100	30	100	6.9	28	150
112	Sorghum	Grain	1600	23	60	6.4	18	100
113	Millet	Grain	1300	20	50	6.1	15	90
114	Lentils	Legume	950	25	70	6.6	22	120
115	Chickpeas	Legume	1050	27	80	6.8	25	130
116	Rapeseed	Oilseed	850	22	60	6.3	20	110
117	Flaxseed	Oilseed	800	20	50	6.0	18	100
118	Coffee	Beverage	2000	25	120	6.5	35	180
119	Tea	Beverage	1800	22	100	6.2	30	160
120	Cocoa	Beverage	1500	28	110	6.8	40	200
121	Avocado	Fruit	1600	30	90	6.7	28	150

5. Data: Astronomy

Task:

Analyze the data below, return your findings in PDF.

Object ID	Object Name	Type	Distance (light years)	Diameter (km)	Mass (solar mass)	Apparent Magnitude
101	Andromeda	Galaxy	2.537 million	110,000	1.230 trillion	3.44
102	Orion Nebula	Nebula	1,344	-	-	4.0
103	Betelgeuse	Star	643	-	11.6	0.50
104	Crab Nebula	Nebula	6,523	-	-	8.4
105	Jupiter	Planet	588 million	139,820	1.898 x 10 ²⁷	-2.94
106	Andromeda II	Galaxy	2.537 million	-	-	14.2
107	Proxima Centauri	Star	4.244	-	0.12	11.13
108	Whirlpool Galaxy	Galaxy	23 million	-	-	8.4
109	Sun	Star	8.3 light minutes	1,391,000	1.989 x 10 ³⁰	-26.74
110	Helix Nebula	Nebula	700	-	-	7.3
111	Saturn	Planet	1.2 billion	116,460	5.683 x 10 ²⁶	0.46
112	Milky Way	Galaxy	-	-	-	-

Object ID	Object Name	Type	Distance (light years)	Diameter (km)	Mass (solar mass)	Apparent Magnitude
113	Pleiades	Star Cluster	440	-	-	1.6
114	Horsehead Nebula	Nebula	1,500	-	-	6.4
115	Alpha Centauri A	Star	4.37	-	1.1	0.01
116	Venus	Planet	0.674 billion	12,104	4.867×10^{24}	-4.92
117	Eagle Nebula	Nebula	7,000	-	-	6.0
118	Polaris	Star	431	-	5.4	2.02
119	Uranus	Planet	2.9 billion	50,724	8.681×10^{25}	5.44
120	Sombrero Galaxy	Galaxy	28 million	-	-	8.99
121	Sirius	Star	8.60	-	2.02	-1.46