

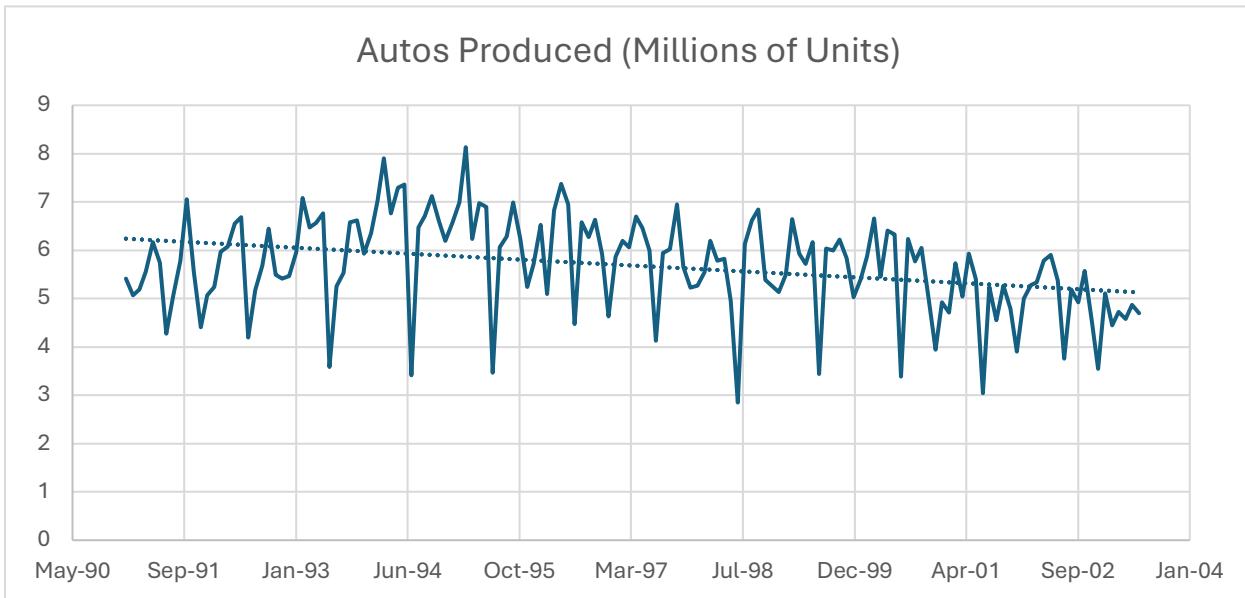
IE 5390 – Assignment - 02

Name: Simran Abhay Sinha

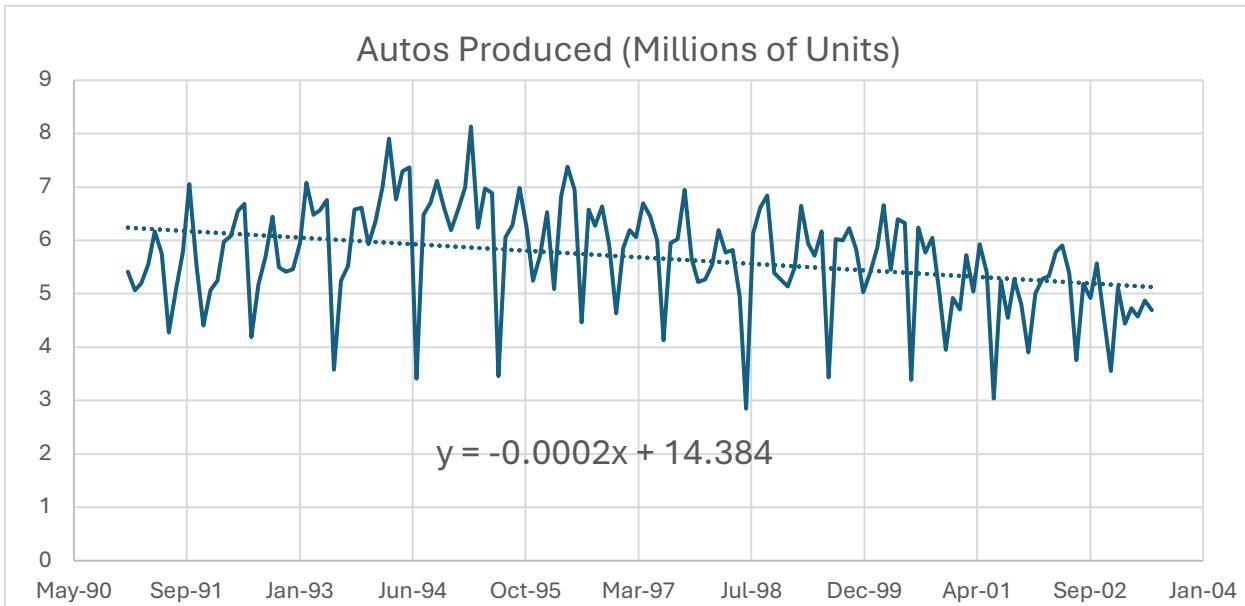
Lab 03

1. Using AutoProduction sheet

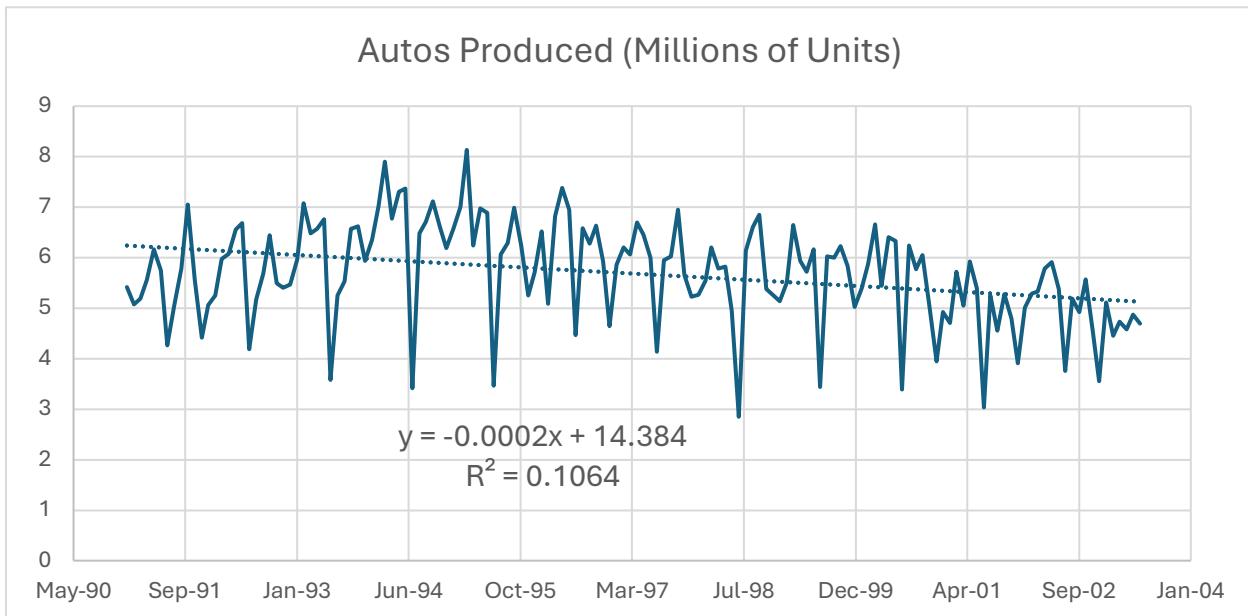
A. Graph the time series of auto production



B. Find the equation of the linear regression of the time series



C. Find the correlation coefficient, R2. What does it mean and what does it say?

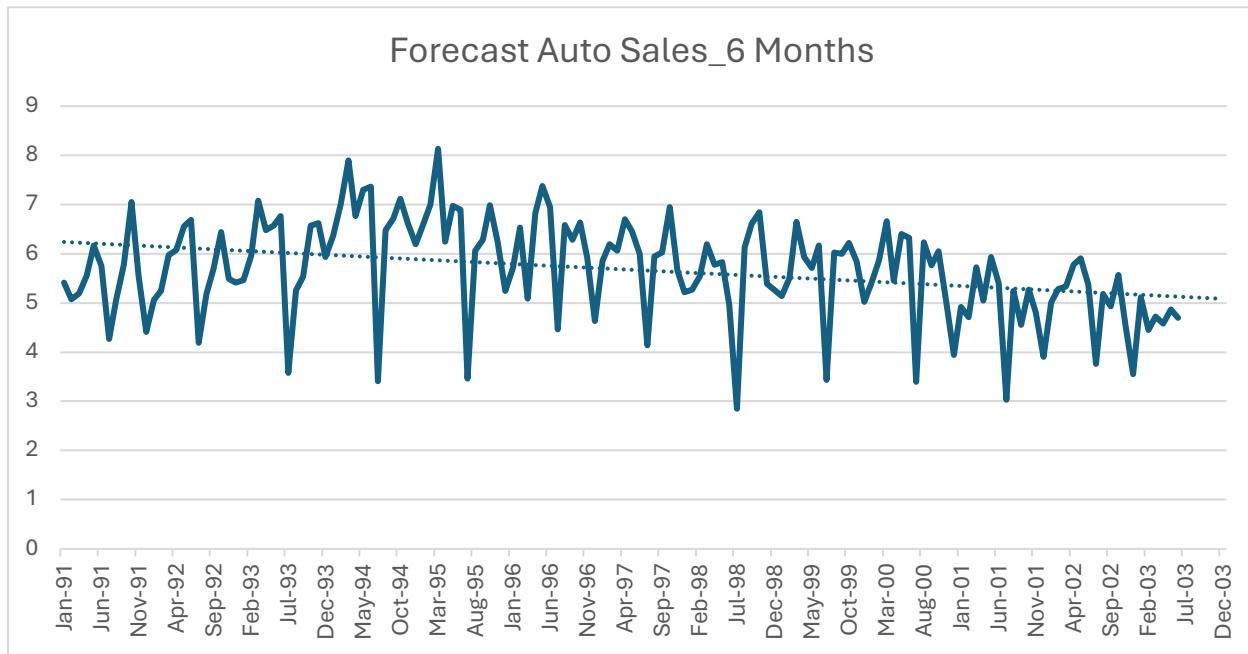


R2 score is displayed in the above graph.

Correlation coefficient R2 represents how well the data fits the linear regression line or curve. For a perfect fit $R^2=1$ and for no fit $R^2=0$. A low R2 value indicates that the trendline does not fit the data well. The above R2 is 0.1064 which means it is a weak fit and the trendline does not fit the data properly.

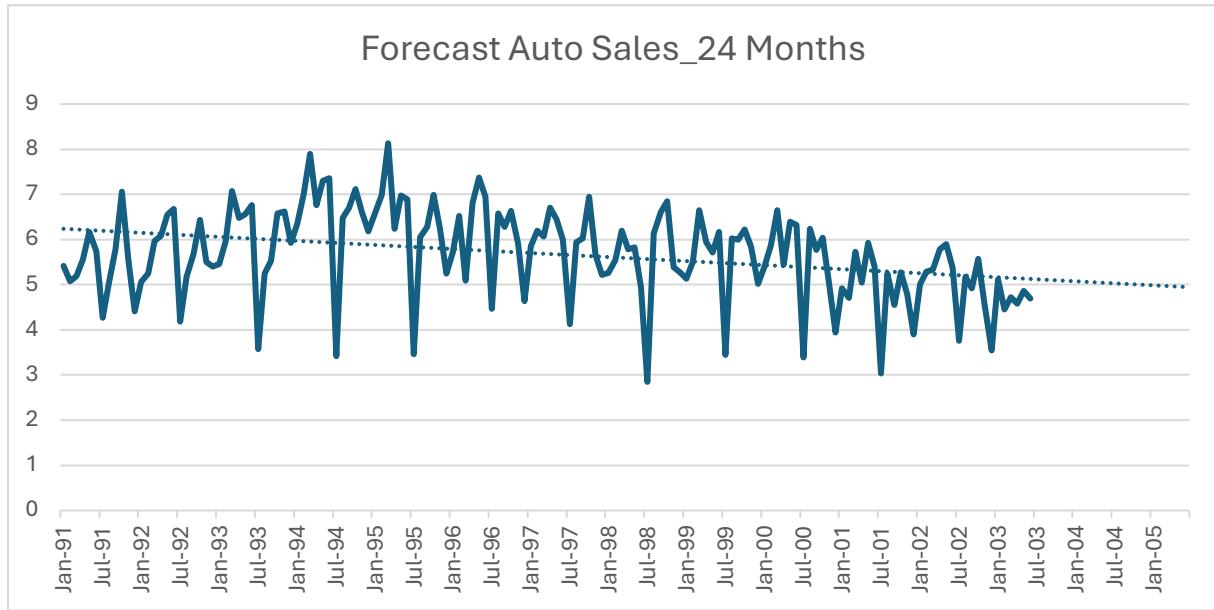
D. Using the graph, forecast the auto sales for the next 6 month

0	May-03	4.869
1	Jun-03	4.694
2	Jul-03	4.650662788
3	Aug-03	3.5483235
4	Sep-03	4.063891578
5	Oct-03	4.72372154
6	Nov-03	5.549984142
7	Dec-03	4.835560789
8		



E. Repeat D, but for the next 24 month

	Index	Date	Value	Forecast
150	150		4.869	
151	151	Jun-03	4.694	Forecast
152	152	Jul-03		4.650662788
153	153	Aug-03		3.5483235
154	154	Sep-03		4.063891578
155	155	Oct-03		4.72372154
156	156	Nov-03		5.549984142
157	157	Dec-03		4.835560789
158	158	Jan-04		3.273939795
159	159	Feb-04		4.223299925
160	160	Mar-04		4.300918473
161	161	Apr-04		4.982086028
162	162	May-04		4.965078191
163	163	Jun-04		4.869853966
164	164	Jul-04		4.239726498
165	165	Aug-04		2.329948642
166	166	Sep-04		4.54333193
167	167	Oct-04		4.686776558
168	168	Nov-04		5.097953356
169	169	Dec-04		3.98374582
170	170	Jan-05		3.016368716
171	171	Feb-05		3.385114963
172	172	Mar-05		4.174136156
173	173	Apr-05		4.897603404
174	174	May-05		4.522727991
175	175	Jun-05		4.669042994
176	176	Jul-05		4.142439595
177	177			
178	178			

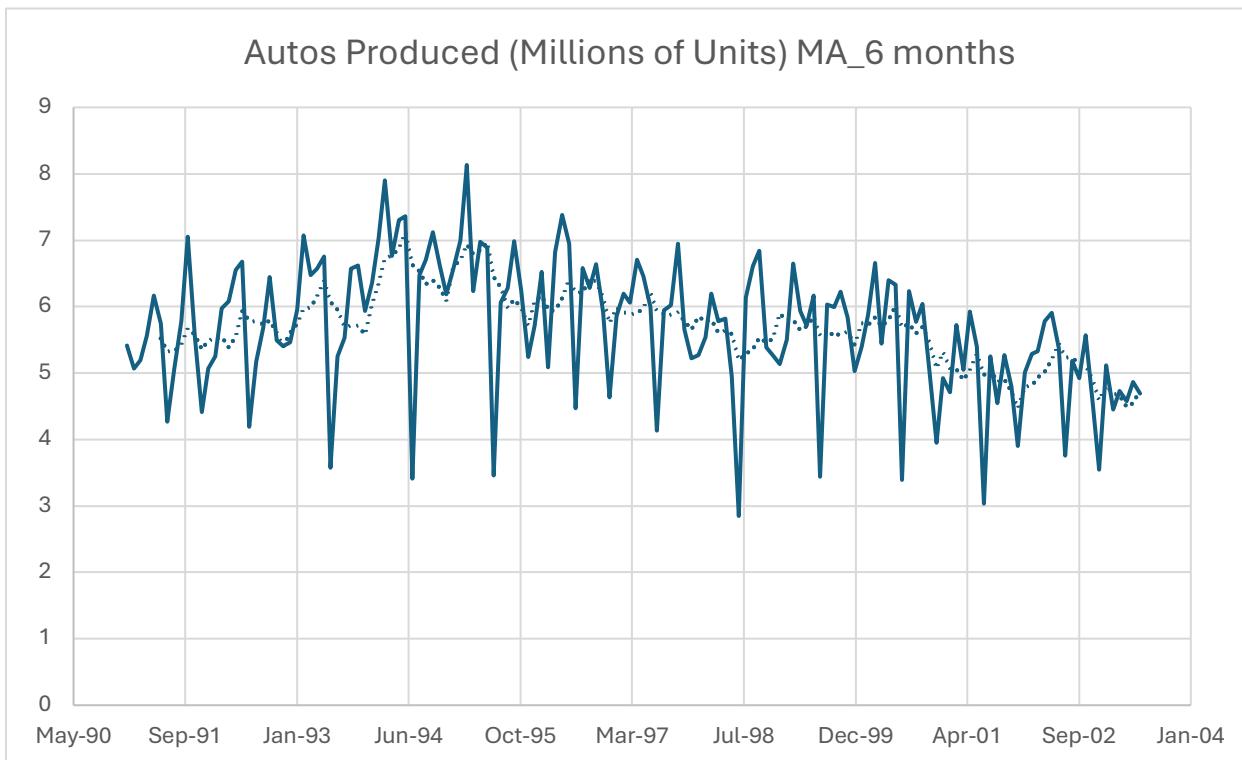


F. Do you trust the forecast results in E. Why or why not?

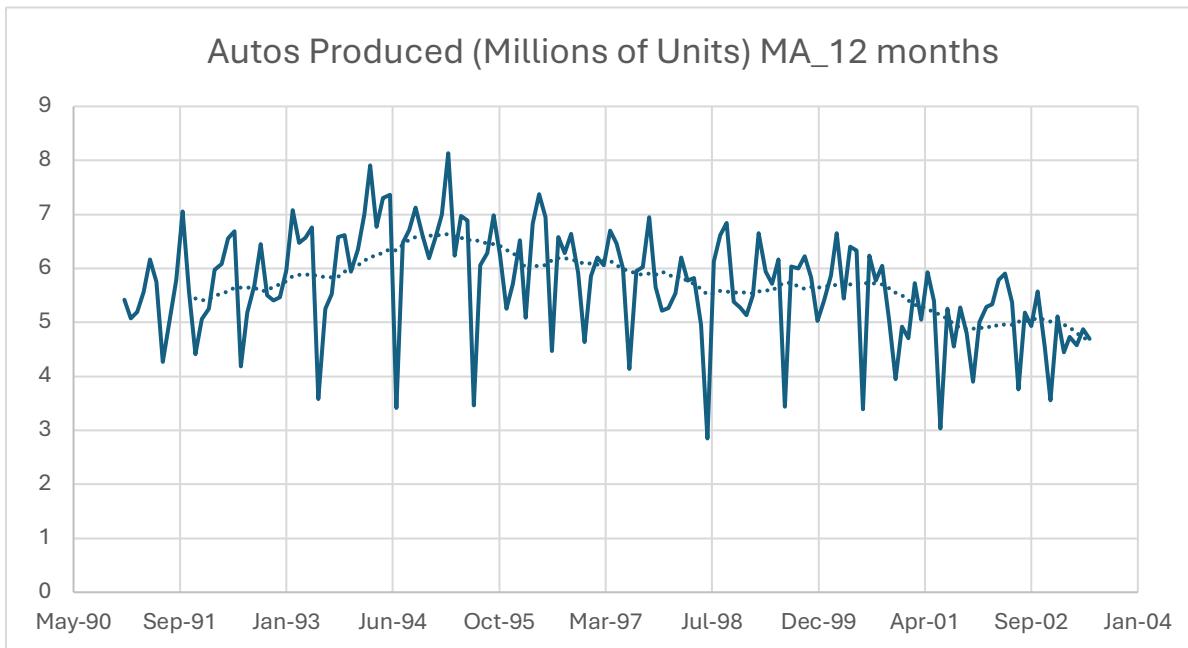
- We cannot trust the forecast results in E as the original data shows fluctuations throughout the year. While, the predicted line is linear which does not fit all the values properly.
- Also, The R2 value is only 0.1064 indicates that the low variation in the data.

2. Using AutoProduction sheet

A. Calculate the 6 month MA



B. Calculate the 12 month MA



C. Which MA provide a better trend? Why

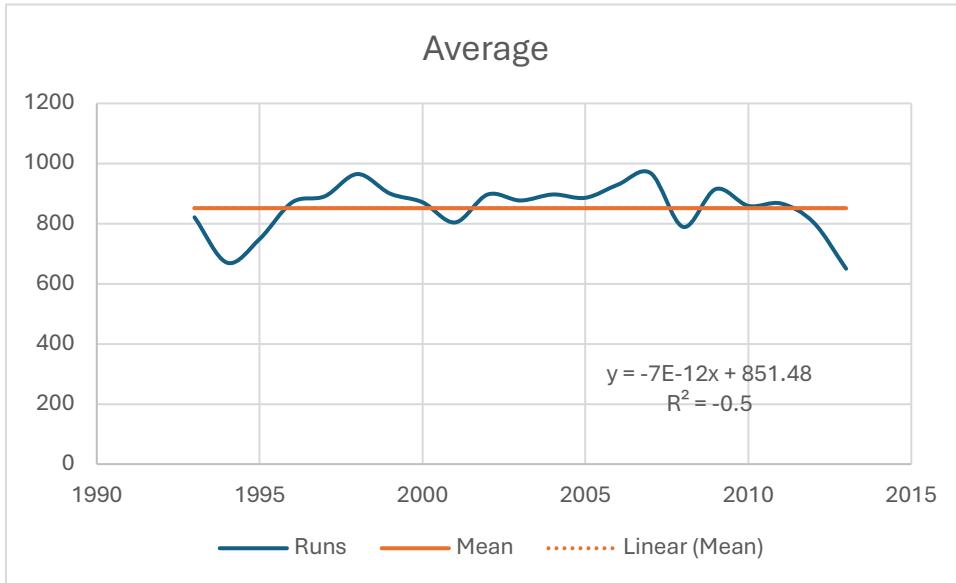
- Auto production has yearly seasonal patterns
- The 12 month MA removes seasonal fluctuations as it is full season cycle.
- The 6 month MA still does not vary that much as it covers only half a year.
- The 6 month MA has more fluctuation which makes trend less clear.

3. Using the YankeeRuns

A. Calculate the average runs for all the years. then copy it in column C

	A	B	C	D	E	F	G
1	Year	Runs	Mean				
2	2013	650	851.48				
3	2012	804	851.48				
4	2011	867	851.48				
5	2010	859	851.48				
6	2009	915	851.48				
7	2008	789	851.48				
8	2007	968	851.48				
9	2006	930	851.48				
10	2005	886	851.48				
11	2004	897	851.48				
12	2003	877	851.48				
13	2002	897	851.48				
14	2001	804	851.48				
15	2000	871	851.48				
16	1999	900	851.48				
17	1998	965	851.48				
18	1997	891	851.48				
19	1996	871	851.48				
20	1995	749	851.48				
21	1994	670	851.48				
22	1993	821	851.48				

B. Graph the time series of the yearly runs and the average to visualize it

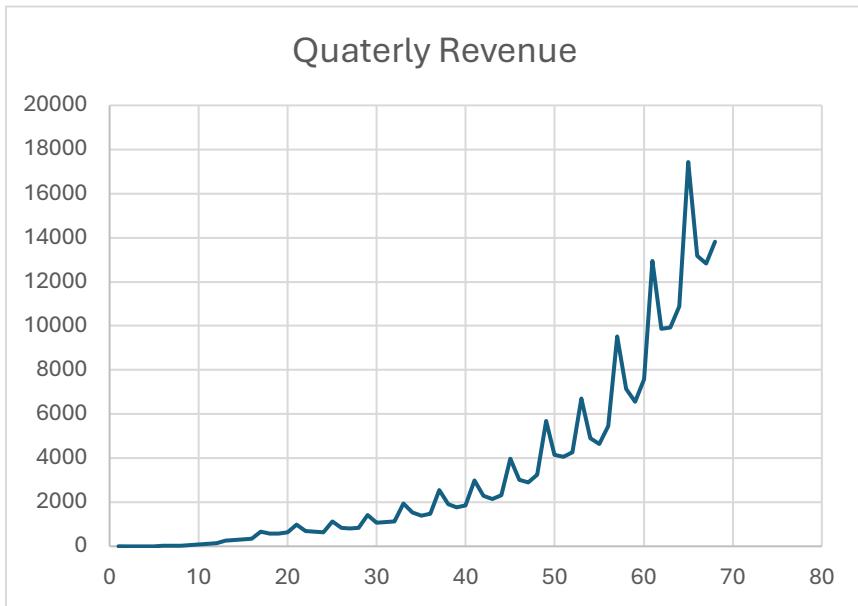


C. What do you see?

- The line graph illustrates runs over time along with its mean.
- The blue line represents runs over the years (1990-2015). It fluctuates showing peaks and troughs. The highest runs are near 2005 and there is a decline near 2015.
- The orange line represents the average value of runs across the period. There is a constant horizontal line around 851.48.
- The dotted line represents the linear trend of the mean which is a flat slope indicating no change in mean over time.

4. Using AmazonRevenues sheet

A. Graph the quarterly revenues



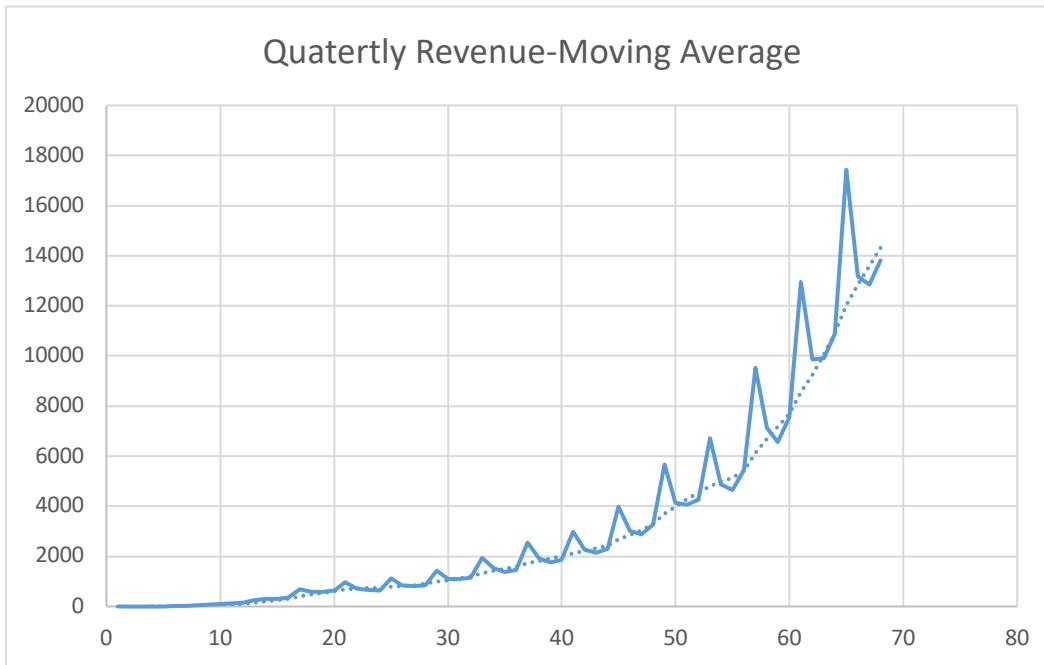
B. Identify the seasonality in the time series

- The above line chart shows the quarterly revenue.
- There are repeating patterns of peaks and troughs throughout the period.
- The peaks indicate the quarter (Q4) where there is a spike due to holiday shopping. There is a more revenue in Q4 as compared to other quarters.

C. Explain why this seasonality occurs

- Seasonality occurs due to sales offered by various brands for shopping events like Christmas, Thanksgiving, Cyber Monday etc.
- Customers generally spend a lot during this season due to promotional offers and events.
- Whereas, Q1 shows lower revenue due to less consumers spending.

D. Create the quarterly MA

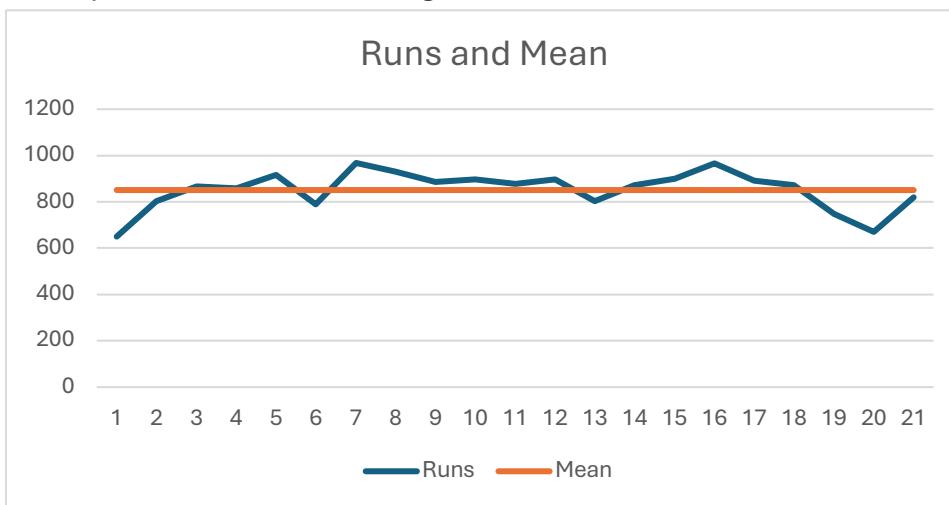


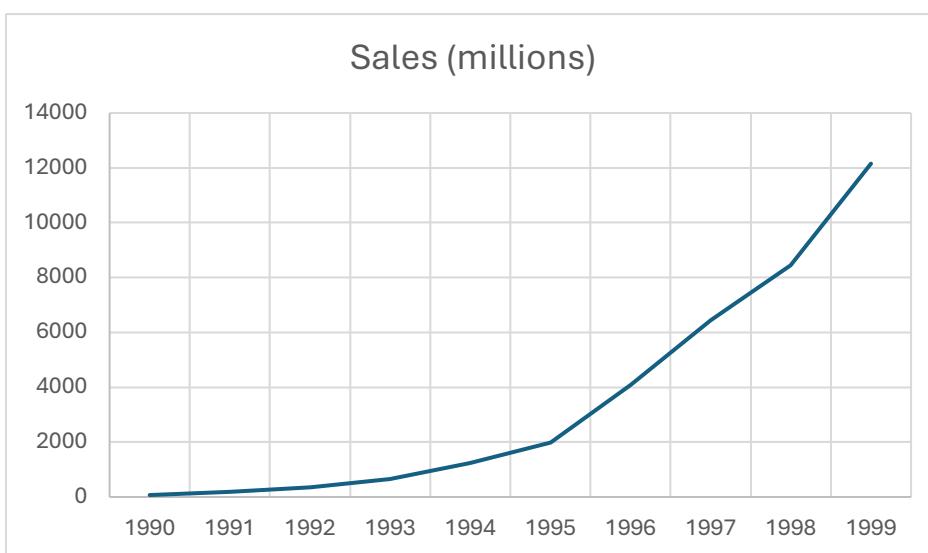
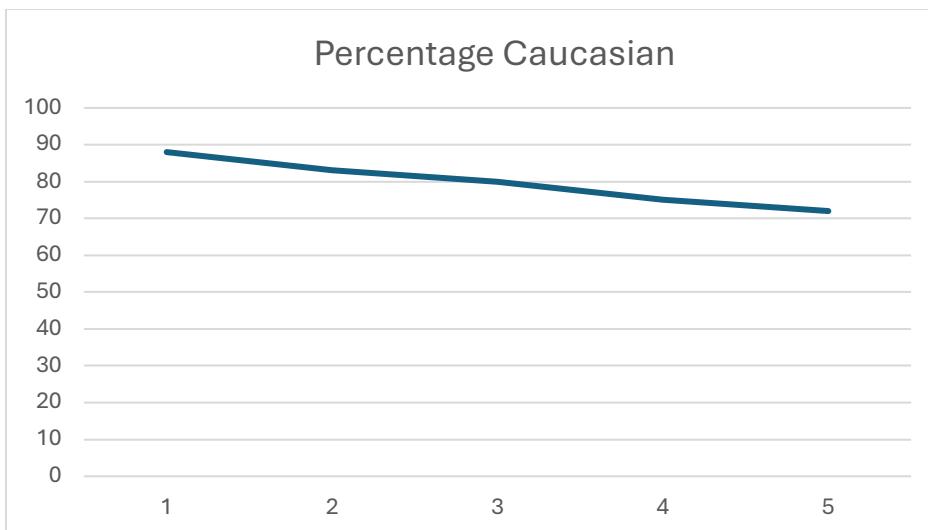
E. What does this MA do?

- The moving average smooths out fluctuation and highlights the overall trend of the data.
- It makes it easier to identify the direction of revenue growth.
- The moving average provides an exhaust view of how companies revenue is changing over time.

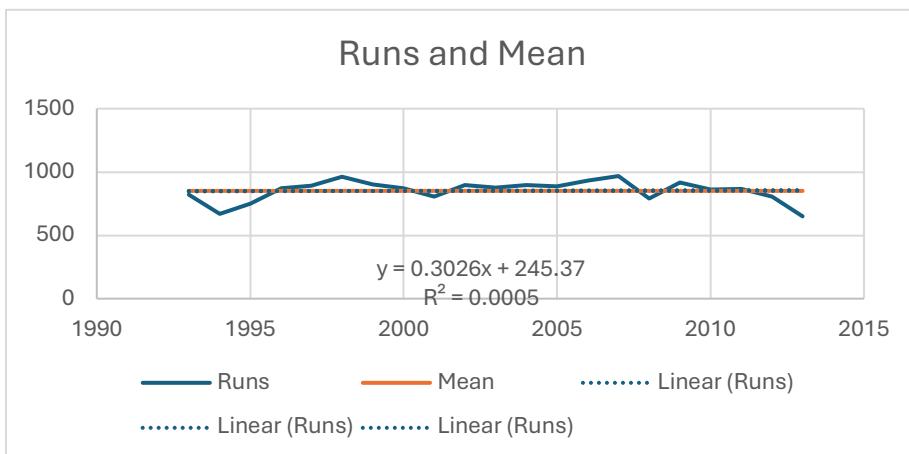
5. Using Three sheet

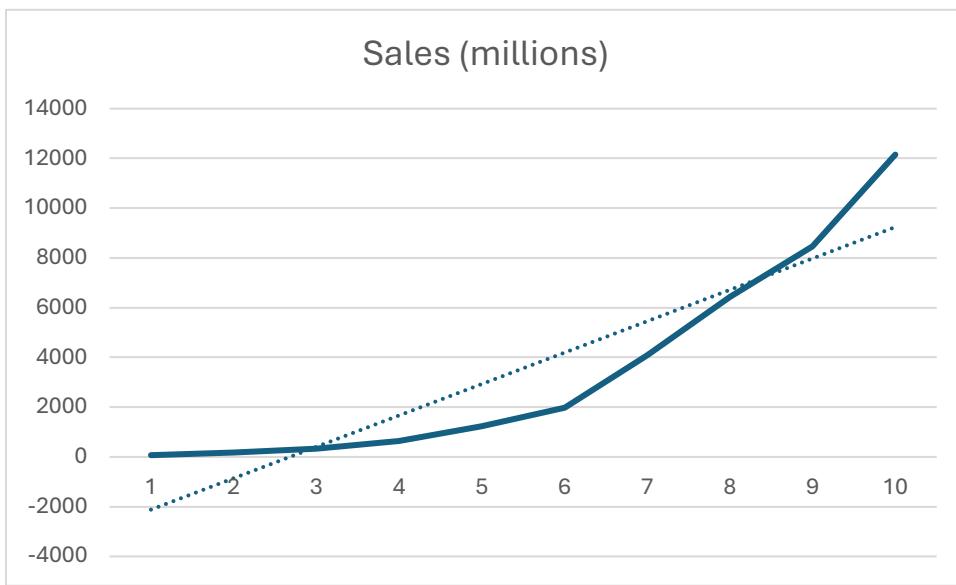
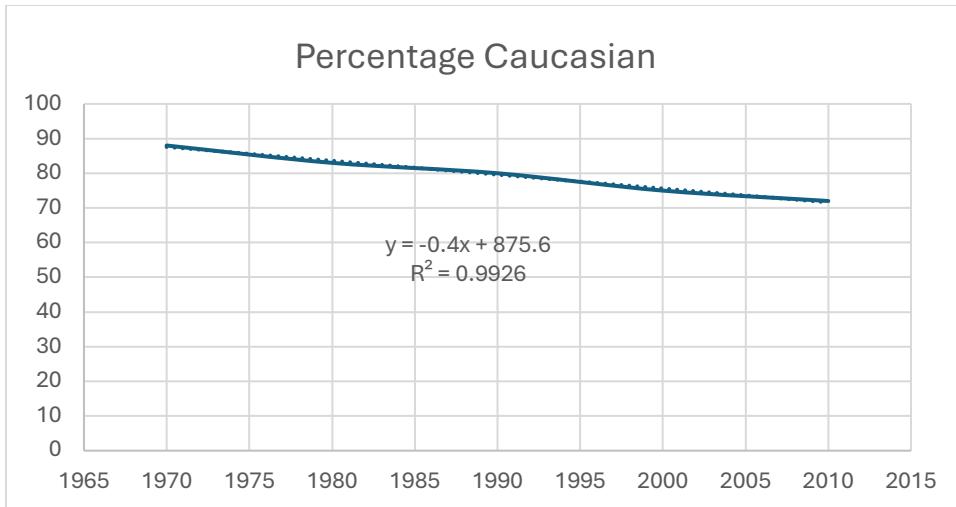
A. Graph the three time series given





B. Insert a trendline for each





C. Which time series have more noise than the other and why?

- Among 3 graphs, “Runs and Mean” has more noise due to many fluctuations and a lack of a clear trend.
- The “Percentage Caucasian” and “Sales” has less noise with smooth trend.

6. Using ThreeF sheet

A. Calculate the error columns shown

fx = \$D3-E3

B	C	D	E	F	G	H	I
Error or Residual = Actual - Forecast							
Month	Sales	Forecast 1	Forecast 2	Forecast 3	Forecast 4		
1	10	12	8	5	15		
2	20	18	22	25	25		
3	15	13	17	10	20		
4	25	27	23	30	30		
5	35	33	37	30	40		
6	40	42	38	45	45		
Month	Error 1	Error 2	Error 3	Error 4			
1	-2	2	5	-5			
2	2	-2	-5	-5			
3	2	-2	5	-5			
4	-2	2	-5	-5			
5	2	-2	5	-5			
6	-2	2	-5	-5			

B. Which forecast is better than the other? Why?

- Forecast 1 and Forecast 2 are more accurate with small error of (+2/-2).
- Forecast 3 and 4 has large error of (+5/-5) units.
- For forecasting sales Forecast 1 or 2 would be better because it has less error thus giving us accurate prediction.

7. Using E1 sheet

A. Calculate MAD

Step1. Calculate the Error

F3 \downarrow \times \checkmark fx =D3-E3

	A	B	C	D	E	F
1	MAD=Mean Absolute Deviation				Average absolute value of errors	
2		Month	Sales	Forecast	Error	
3		1	37	26	11	
4		2	32	40	-8	
5		3	28	45	-17	
6		4	32	38	-6	
7		5	32	22	10	
8		6	32	22	10	

Step 2: Removing Absolute Error

\downarrow \times \checkmark fx =ABS(F3)

	B	C	D	E	F	G	H	I	J
	MAD=Mean Absolute Deviation				Average absolute value of errors				
		Month	Sales	Forecast	Error	Absolute			
		1	37	26	11	11			
		2	32	40	-8	8			
		3	28	45	-17	17			
		4	32	38	-6	6			
		5	32	22	10	10			
		6	32	22	10	10			

Step 3: Take the average of Absolute Error

B	C	D	E	F	G	H	I	J
$=AVERAGE(G3:G8)$								
MAD=Mean Absolute Deviation					Average absolute value of errors			
Month	Sales	Forecast	Error	Absolute Error				
1	37	26	11	11				
2	32	40	-8	8				
3	28	45	-17	17				
4	32	38	-6	6				
5	32	22	10	10				
6	32	22	10	10				
Average Error								
MAD					10.333333			

B. Is this forecasting model good or bad? Explain your answer

- The forecasting model is bad because it has high MAD (Mean Absolute Deviation) of 10.3333. This indicates that there will be large errors in predictions.
- The average error of 0 is misleading because it is just canceling out the positive and negative numbers.
- In this case MAD is a valuable metric to choose between the forecasting models.

8. Using E2 sheet

A. Calculate MAD

Step1. Calculate the Error

F5	A	B	C	D	E	F
	MAPE=Mean Absolute Percentage Error Percentage Error=(Actual-forecast)/Actual					
1						
Series 1						
4		Month	Sales	Forecast	Error	
5		1	20	23	-3	
6		2	30	27	3	
7		3	40	43	-3	
8		4	50	47	3	
9						
10						
11		Series 2				
12		Month	Sales	Forecast	Error	
13		1	200	210	-10	
14		2	300	290	10	
15		3	400	390	10	
16		4	500	510	-10	

Step 2: Removing Absolute Error

F5	A	B	C	D	E	F	G
	MAPE=Mean Absolute Percentage Error Percentage Error=(Actual-forecast)/Actual						
1							
Series 1							
	Month	Sales	Forecast	Error	Absolute error		
	1	20	23	-3	3		
	2	30	27	3	3		
	3	40	43	-3	3		
	4	50	47	3	3		
Series 2							
	Month	Sales	Forecast	Error	Absolute error		
	1	200	210	-10	10		
	2	300	290	10	10		
	3	400	390	10	10		
	4	500	510	-10	10		

Step 3: Average of Absolute Error

A	B	C	D	E	F	G	H
$f_x = \text{AVERAGE}(G5:G8)$							
MAPE=Mean Absolute Percentage Error Percentage Error=(Actual-forecast)/Actual							
						MAD	
			Series 1			3	
		Month	Sales	Forecast	Error	Absolute error	Percentage Error
		1	20	23	-3	3	-0.15
		2	30	27	3	3	0.1
		3	40	43	-3	3	-0.075
		4	50	47	3	3	0.06
						MAD	
			Series 2			10	
		Month	Sales	Forecast	Error	Absolute error	Percentage Error
		1	200	210	-10	10	-0.05
		2	300	290	10	10	0.0333
		3	400	390	10	10	0.025
		4	500	510	-10	10	-0.02

B. Calculate MAPE

Step 1: Calculating APE using percentage error

A	B	C	D	E	F	G	H
Absolute Percentage Error APE=(Actual-forecast)/Actual							
MAPE							
			Series 1			3	0.09625
		Month	Sales	Forecast	Error	Absolute error	Percentage Error
		1	20	23	-3	3	-0.15
		2	30	27	-3	3	-0.1
		3	40	43	-3	3	-0.075
		4	50	47	3	3	0.06

					MAD	MAPE
Series 2					10	0.03208
Month	Sales	Forecast	Error	Absolute error	Percentage Error	APE
1	200	210	-10	10	-0.05	0.05
2	300	290	10	10	0.0333	0.0333
3	400	390	10	10	0.025	0.025
4	500	510	-10	10	-0.02	0.02

Step 2: Removing average of APE to find MAPE

Absolute Percentage Error APE = $\frac{\text{Actual} - \text{Forecast}}{\text{Actual}}$					MAD	MAPE
Series 1					3	0.09625
Month	Sales	Forecast	Error	Absolute error	Percentage Error	APE
1	20	23	-3	3	-0.15	0.15
2	30	27	-3	3	-0.1	0.1
3	40	43	-3	3	-0.075	0.075
4	50	47	3	3	0.06	0.06

Series 2					MAD	MAPE
Series 2					10	0.03208
Month	Sales	Forecast	Error	Absolute error	Percentage Error	APE
1	200	210	-10	10	-0.05	0.05
2	300	290	10	10	0.0333	0.0333
3	400	390	10	10	0.025	0.025
4	500	510	-10	10	-0.02	0.02

C. What is your conclusion?

- After analyzing series 1 and series 2, Series 1 has a low MAD compared to Series 2. But, the MAPE of Series 2 is 3.2% and that of Series 1 is 9.6% indicating more accuracy.
- The absolute error in Series 2 is larger because the data values are also larger. But, the MAPE is 3.2% which indicates it is performing better overall.

9. Using E3 sheet, calculate SSE

Step 1: Calculating Error

A	B	C	D	E	F
SSE=Sum of Squared forecast errors					
Series 1					
	Month	Sales	Forecast	Error	
	1	20	23	-3	
	2	30	27	3	
	3	40	43	-3	
	4	50	46	4	
Series 2					
	Month	Sales	Forecast	Error	
	1	200	210	-10	
	2	300	290	10	
	3	400	390	10	
	4	500	510	-10	

Step 2: Removing Squared Error

G4	A	B	C	D	E	F	G
1	SSE=Sum of Squared forecast errors						SSE
2	Series 1						43
3		Month	Sales	Forecast	Error	Squared Error	
4		1	20	23	-3	9	
5		2	30	27	3	9	
6		3	40	43	-3	9	
7		4	50	46	4	16	
8	Series 2						SSE
9		Month	Sales	Forecast	Error	Squared Error	400
10		1	200	210	-10	100	
11		2	300	290	10	100	
12		3	400	390	10	100	
13		4	500	510	-10	100	

Step 3: Calculating SSE

G2	A	B	C	D	E	F	G	H
1	SSE=Sum of Squared forecast errors					SSE		
2	Series 1						43	
3		Month	Sales	Forecast	Error	Squared Error		
4		1	20	23	-3	9		
5		2	30	27	3	9		
6		3	40	43	-3	9		
7		4	50	46	4	16		
8								
9						SSE		
10		Series 2					400	
11		Month	Sales	Forecast	Error	Squared Error		
12		1	200	210	-10	100		
13		2	300	290	10	100		
14		3	400	390	10	100		
15		4	500	510	-10	100		
16								

10. Using E4 sheet

A. Calculate the bias

Step 1: Calculating Actual/Forecast-1

B	C	D	E	F
Month	Sales	Forecast 1	Actual/Forecast -1	
1	10	8	0.25	
2	20	25	-0.2	
3	15	10	0.5	
4	25	30	-0.166666667	
5	35	30	0.166666667	
Month	Sales	Forecast 2	Actual/Forecast -1	
1	10	15	-0.333333333	
2	20	25	-0.2	
3	15	20	-0.25	
4	25	30	-0.166666667	
5	35	40	-0.125	

Step 2: Removing Average of Actual/Forecast-1

	A	B	C	D	E	F	G	H	I	J
1								Mean bias Forecast 1		
2		Month	Sales	Forecast 1	Actual/Forecast -1			Mean bias Forecast 2		
3		1	10	8	0.25					
4		2	20	25	-0.2					
5		3	15	10	0.5					
6		4	25	30	-0.166666667					
7		5	35	30	0.166666667					
8										
9										
10		Month	Sales	Forecast 2	Actual/Forecast -1					
11		1	10	15	-0.333333333					
12		2	20	25	-0.2					
13		3	15	20	-0.25					
14		4	25	30	-0.166666667					
15		5	35	40	-0.125					

B. What is your conclusion?

- Forecast 1 has a mean bias of 0.11 whereas, forecast 2 has mean bias of -0.215 which indicates that Forecast 1 overestimates the actual values on average and Forecast 2 underestimates it.
 - Forecast 1 performs better as the bias is closer to 0 and it would be a better choice for future predictions.

11. Using E5, determine if the Forecast 1 bias is significant and if Forecast 2 bias is significant

Step 1: Calculating Std Dev Bias

B	C	D	E	F	G	H	I	J	K	L
					11.00%	Mean bias Forecast 1		Std dev Bias Forecast 1	29.48%	
Month	Sales	Forecast 1	Actual/Forecast -1		-21.50%	Mean bias Forecast 2		Std dev Bias Forecast 2	8.05%	
1	10	8		0.25			insignificant bias	significant bias		
2	20	25		-0.2		95% Confidence Interval for Bias	Forecast 1		Forecast 2	
3	15	10		0.5		Mean bias-2*Stdevbias/observations		-26.36%		-7.20%
4	25	30		-0.166666667		Mean bias+2*Stdevbias/observations		26.36%		7.20%
							O in 95% Confidence Interval:Bias not significant			
5	35	30	0.166666667			O not in 95% Confidence Interval:Bias is significant				
							O in 95% Confidence Interval:Bias not significant			
Month	Sales	Forecast 2	Actual/Forecast -1							
1	10	15		-0.333333333						
2	20	25		-0.2						
3	15	20		-0.25						
4	25	30		-0.166666667						
5	35	40		-0.125						

Step 2: Calculating Insignificant and Insignificant Bias

B	C	D	E	F	G	H	I	J	K	L
					11.00% Mean bias Forecast 1			Std dev Bias Forecast 1	29.48%	
Month	Sales	Forecast 1	Actual/Forecast -1	-0.25	-21.50% Mean bias Forecast 2			Std dev Bias Forecast 2	8.05%	
1	10	8		0.25	95% Confidence Interval for Bias		insignificant bias	Forecast 1		significant bias
2	20	25		-0.2	Mean bias-2*Stdevbias/Vobservations				-26.36%	-7.20%
3	15	10		0.5	Mean bias+2*Stdevbias/Vobservations				26.36%	7.20%
4	25	30		-0.166666667						
							0 in 95% Confidence Interval: Bias is significant			
							0 not in 95% Confidence Interval: Bias not significant			
5	35	30		0.166666667			0 in 95% Confidence Interval: Bias not significant			
Month	Sales	Forecast 2	Actual/Forecast -1							
1	10	15		-0.333333333						
2	20	25		-0.2						
3	15	20		-0.25						
4	25	30		-0.166666667						
5	35	40		-0.125						

C	D	E	F	G	H	I	J	K	L	M
				11.00% Mean bias Forecast 1			Std dev Bias Forecast 1	29.48%		
Month	Sales	Forecast 1	Actual/Forecast -1	-21.50% Mean bias Forecast 2			Std dev Bias Forecast 2	8.05%		
1	10	8		0.25	95% Confidence Interval for Bias		insignificant bias	Forecast 1		significant bias
2	20	25		-0.2	Mean bias-2*Stdevbias/Vobservations			Forecast 2		
3	15	10		0.5	Mean bias+2*Stdevbias/Vobservations				-26.36%	-7.20%
4	25	30		-0.166666667	Mean bias-2*Stdevbias/Vobservations				26.36%	7.20%
							0 in 95% Confidence Interval: Bias is significant			
							0 not in 95% Confidence Interval: Bias not significant			
5	35	30		0.166666667			0 in 95% Confidence Interval: Bias not significant			
Month	Sales	Forecast 2	Actual/Forecast -1							
1	10	15		-0.333333333						
2	20	25		-0.2						
3	15	20		-0.25						
4	25	30		-0.166666667						
5	35	40		-0.125						

12. **Gen AI Prompt Engineering:** use your favorite AI Chat Bot to solve the following problem:

I am using Perplexity.

A. Use the excel file with the name Temperature_Forecasting.xlsx. it has 15 days of previous temps.

B. Feed the data to Gen AI bot and ask it to use it to create a linear regression model
Prompt:

Generate a line regression model for the above data in excel and give the graph

Gen AI:



To generate a line regression model for the given temperature data and create a graph in Excel, follow these steps:

1. Data Preparation:
The data is already organized in a suitable format with dates and corresponding temperatures.
2. Create Scatter Plot:
 - Select the Date and Temperature columns.
 - Insert a Scatter chart.
 - Choose the option for a scatter plot with smooth lines and markers.
3. Add Trendline:
 - Right-click on any data point in the chart.
 - Select "Add Trendline" from the context menu.
 - Choose "Linear" as the trendline type.
 - Check the boxes for "Display Equation on chart" and "Display R-squared value on chart".
4. Formatting:
 - Add appropriate titles for the chart and axes.
 - Adjust the chart area and plot area as needed for clarity.

Here's a description of the resulting graph:

Temperature Trend Analysis (January 1-14, 2025)

The graph displays a scatter plot of daily temperatures from January 1 to January 14, 2025, with a linear regression line. The x-axis represents the date, and the y-axis shows the temperature in degrees Celsius.

Key Features:

- Data points are represented by blue markers.
- A red linear trendline runs through the data points.
- The trendline equation is displayed on the chart: $y = 0.3x + 14.371$
- The R-squared value is also shown: $R^2 = 0.9441$

Interpretation:

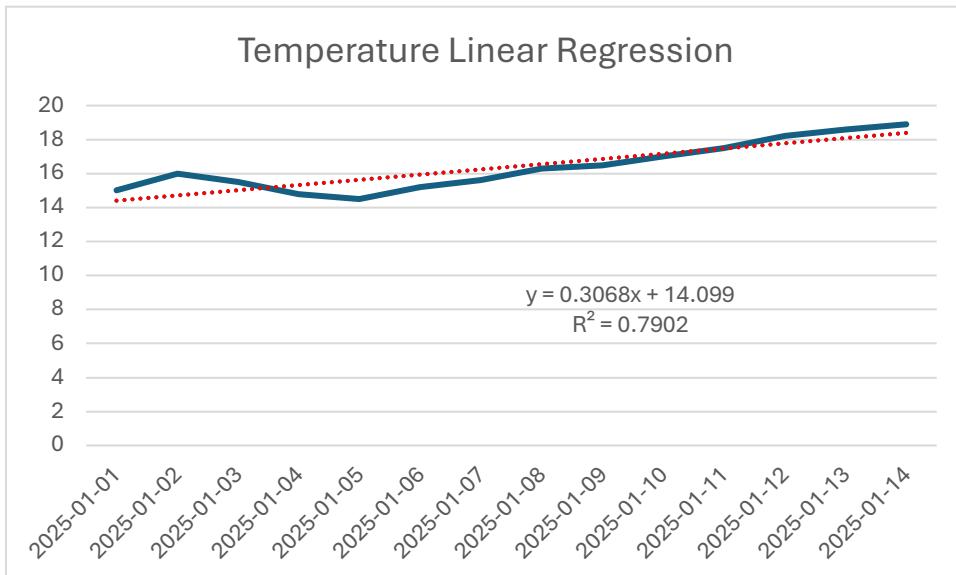
- The positive slope (0.3) indicates an upward trend in temperature over the two-week period.
- The R-squared value of 0.9441 suggests a strong fit of the linear model to the data, with about 94.41% of the temperature variation explained by the date.
- The y-intercept (14.371) represents the theoretical starting temperature on January 1, 2025.

This linear regression model can be used to predict temperatures for dates within or slightly beyond the given range, assuming the trend continues .

Gen AI was not able to generate the graph instead, it gave the description of graph.

C. Redo Steps B and C above in excel

I performed the steps in Excel and generated the graph.



D. Compare excel model with Gen AI model: compare the coefficients of the two regression models, i.e. compare the Y-intercept and the slope, say β_0 and β_1 in the equation: $Y = \beta_0 + \beta_1 X + \varepsilon$. Are they equal? They should be. If not try other bots.

Excel:

- Y interpret= 14.099
- Slope: 0.3068
- R2= 0.7902

Gen AI:

- Y interpret= 14.371
- Slope: 0.3
- R2= 0.9441

Comparison of Excel and Gen AI:

- Y interpret : Both models have given similar value.
- Slope: Both model shows the same slope (0.3068)
- R2 values are not same in the both model.

E. Ask Gen AI to predict the temperature for days 16-21

Prompt:

Using the above dataset predict the temperature for days 16-21 and also gives the steps for excel

Gen AI:

Perplexity

To predict the temperature for days 16-21 using the given dataset, we'll use the linear regression model we developed earlier and extend it in Excel. Here are the steps to do this and the predicted temperatures:

Steps in Excel

1. Extend the day numbers:
 - In column E, continue the day numbers from 15 to 21.
2. Extend the date column:
 - In column A, use the formula =A14+1 in cell A15 and drag down to A21.
3. Calculate predicted temperatures:
 - Assuming you've already calculated the slope and intercept as described earlier, in cell F15, use the formula:

```
text
=$G$2*E15 + $H$2
```

- Drag this formula down to F21.

Predicted Temperatures

Using the linear regression equation $Y = 14.371 + 0.3X$, where X is the day number, we can predict the temperatures for days 16-21:

Day	Date	Predicted Temperature (°C)
16	2025-01-16	19.17
17	2025-01-17	19.47
18	2025-01-18	19.77
19	2025-01-19	20.07
20	2025-01-20	20.37
21	2025-01-21	20.67

These predictions suggest a continued warming trend, with temperatures rising by 0.3°C each day. However, it's important to note that these are extrapolations based on the existing data and may not account for potential weather fluctuations or other factors that could influence actual temperatures [1](#) [2](#).

Excel:

Day	Date	Temperature (°C)
1	2025-01-01	15
2	2025-01-02	16
3	2025-01-03	15.5
4	2025-01-04	14.8
5	2025-01-05	14.5
6	2025-01-06	15.2
7	2025-01-07	15.6
8	2025-01-08	16.3
9	2025-01-09	16.5
10	2025-01-10	17
11	2025-01-11	17.5
12	2025-01-12	18.2
13	2025-01-13	18.6
14	2025-01-14	18.9
15	2025-01-15	19.90
16	2025-01-16	20.20
17	2025-01-17	20.30
18	2025-01-18	20.33
19	2025-01-19	20.34
20	2025-01-20	20.34
21	2025-01-21	20.34

F. To measure the accuracy of the model, ask Gen AI bot to calculate the temperatures for days 5 – 15. The bot should be able to use the independent variable x (day # in this problem) and calculate the dependent variable Y (temperature in this problem)

Prompt:

Using the above excel file calculate temperature of day 5-15

Gen AI:

Complete Temperature List (Days 5-15)

Day	Date	Temperature (°C)
5	2025-01-05	14.5
6	2025-01-06	15.2
7	2025-01-07	15.6
8	2025-01-08	16.3
9	2025-01-09	16.5
10	2025-01-10	17.0
11	2025-01-11	17.5
12	2025-01-12	18.2
13	2025-01-13	18.6
14	2025-01-14	18.9
15	2025-01-15	18.9 (predicted)



This list provides the temperatures for days 5-15, with the last day's temperature being a prediction based on the linear regression model derived from the entire dataset.

Perplexity

To calculate the temperature for days 5-15 using the provided Excel file, I'll use the given data for days 5-14 and extend the linear regression model to predict day 15. Here's the calculation:

Actual Temperatures (Days 5-14)

Day	Date	Temperature (°C)
5	2025-01-05	14.5
6	2025-01-06	15.2
7	2025-01-07	15.6
8	2025-01-08	16.3
9	2025-01-09	16.5
10	2025-01-10	17.0
11	2025-01-11	17.5
12	2025-01-12	18.2
13	2025-01-13	18.6
14	2025-01-14	18.9



Predicted Temperature (Day 15)

To predict the temperature for day 15, we need to calculate the slope and intercept of the linear regression line using the entire dataset (days 1-14).

Using Excel functions:

1. Slope (β_1): =SLOPE(C2:C15, A2:A15) ≈ 0.3
2. Intercept (β_0): =INTERCEPT(C2:C15, A2:A15) ≈ 14.371

The linear regression equation is:

$$Y = 14.371 + 0.3X$$

For day 15:

$$Y = 14.371 + 0.3 * 15 \approx 18.871$$

Therefore, the predicted temperature for day 15 (2025-01-15) is approximately 18.9°C.

G. Ask Gen AI to create a table with actual values, predicted values, and the error (difference) for days 5 – 15

Prompt:

Create a table with actual values that are present in the sheet, predicted values and the error difference for days 5-15
table should have 3 columns of actual, predicted and error

Gen AI:

Based on the data provided in the Excel sheet and using the linear regression model, I'll create a table with actual values, predicted values, and the error difference for days 5-15. First, let's calculate the slope and intercept:

Slope (β_1) ≈ 0.3

Intercept (β_0) ≈ 14.371

Using the equation $Y = 14.371 + 0.3X$, where X is the day number, we can calculate the predicted temperatures.

Day	Actual (°C)	Predicted (°C)	Error (°C)
5	14.5	15.871	1.371
6	15.2	16.171	0.971
7	15.6	16.471	0.871
8	16.3	16.771	0.471
9	16.5	17.071	0.571
10	17.0	17.371	0.371
11	17.5	17.671	0.171
12	18.2	17.971	-0.229
13	18.6	18.271	-0.329
14	18.9	18.571	-0.329
15	N/A	18.871	N/A

Day	Actual (°C)	Predicted (°C)	Error (°C)
14	18.9	18.571	-0.329
15	N/A	18.871	N/A

Note: The actual temperature for day 15 is not provided in the original dataset, so we can't calculate the error for that day. The predicted values are rounded to three decimal places for consistency.

H. Ask Gen AI to calculate the model accuracy: ask it which error formula it used

Prompt:

Calculate the model accuracy and mention the error formula which you used

Gen AI:

To calculate the model accuracy for the temperature forecasting data, we'll use the Mean Absolute Error (MAE) as our error metric. The MAE measures the average magnitude of errors in a set of predictions, without considering their direction.

Error Formula

The Mean Absolute Error (MAE) is calculated using the following formula:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Where:

- n is the number of observations
- y_i is the actual value
- \hat{y}_i is the predicted value

Calculation

First, let's calculate the predicted values using our linear regression model:
 $Y = 14.371 + 0.3X$

First, let's calculate the predicted values using our linear regression model:

$$Y = 14.371 + 0.3X$$

Day	Actual (°C)	Predicted (°C)	Absolute Error
1	15.0	14.671	0.329
2	16.0	14.971	1.029
3	15.5	15.271	0.229
4	14.8	15.571	0.771
5	14.5	15.871	1.371
6	15.2	16.171	0.971
7	15.6	16.471	0.871
8	16.3	16.771	0.471
9	16.5	17.071	0.571
10	17.0	17.371	0.371
11	17.5	17.671	0.171
12	18.2	17.971	0.229
13	18.6	18.271	0.329
14	18.9	18.571	0.329

Sum of Absolute Errors: 8.042

MAE = $8.042 / 14 \approx 0.574^{\circ}\text{C}$

Model Accuracy

The Mean Absolute Error of 0.574°C indicates that, on average, our model's predictions deviate from the actual temperatures by about 0.574°C . This suggests a reasonably good fit, considering the natural variability in daily temperatures.

To express this as a percentage accuracy, we can use the following formula:

$$\text{Accuracy} = (1 - \text{MAE} / \text{Range}) * 100$$

Where Range is the difference between the maximum and minimum temperatures in our dataset.

$$\text{Range} = 18.9^{\circ}\text{C} - 14.5^{\circ}\text{C} = 4.4^{\circ}\text{C}$$

$$\text{Accuracy} = (1 - 0.574 / 4.4) * 100 \approx 86.95\%$$

Therefore, our model has an accuracy of approximately 86.95% based on the Mean Absolute Error metric.

- I. Ask Gen AI to create a line chart showing the actual and predicted temperatures for days 5 – 15

Prompt:

can you create a line chart showing the actual and predicted temperatures for days 5 – 15

Gen AI:

✳️ Perplexity

I apologize, but I cannot create an actual line chart or image for you. However, I can describe how to create a line chart in Excel using the given data and the linear regression model we've calculated earlier.

To create a line chart showing the actual and predicted temperatures for days 5-15:

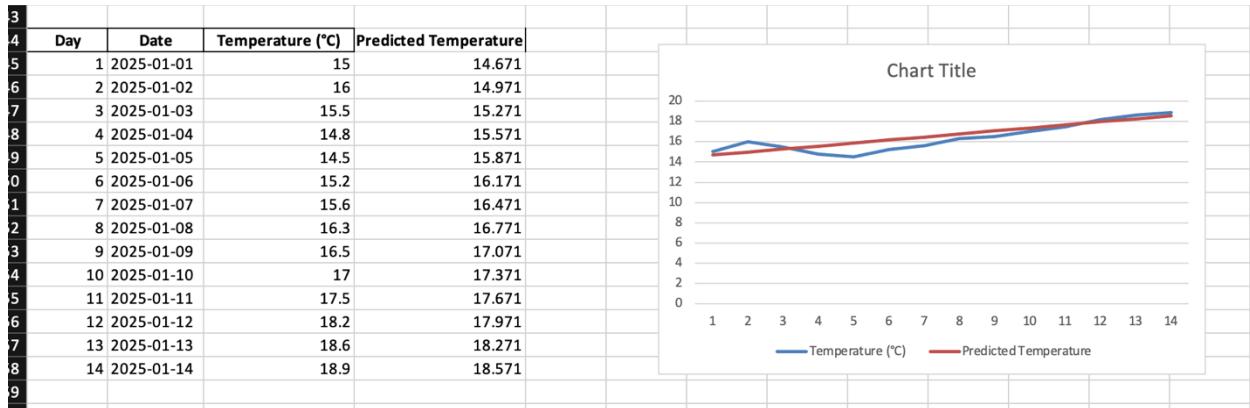
1. In Excel, select columns for Day (A), Actual Temperature (C), and Predicted Temperature (which you'll need to calculate).
2. Calculate predicted temperatures using the formula:
 $\text{Predicted Temperature} = 14.371 + 0.3 * \text{Day Number}$
3. Insert a line chart:
 - Go to the Insert tab
 - Click on "Line" chart
 - Select "Line with Markers"
4. Customize the chart:
 - Add a title: "Actual vs Predicted Temperatures (Days 5-15)"
 - Label the X-axis: "Day"
 - Label the Y-axis: "Temperature ($^{\circ}\text{C}$)"
 - Add a legend to distinguish between actual and predicted lines

The resulting chart would show:

- X-axis: Days 5 through 15
- Y-axis: Temperature scale from approximately 14°C to 19°C

The Gen AI was not able to generate the line chart.

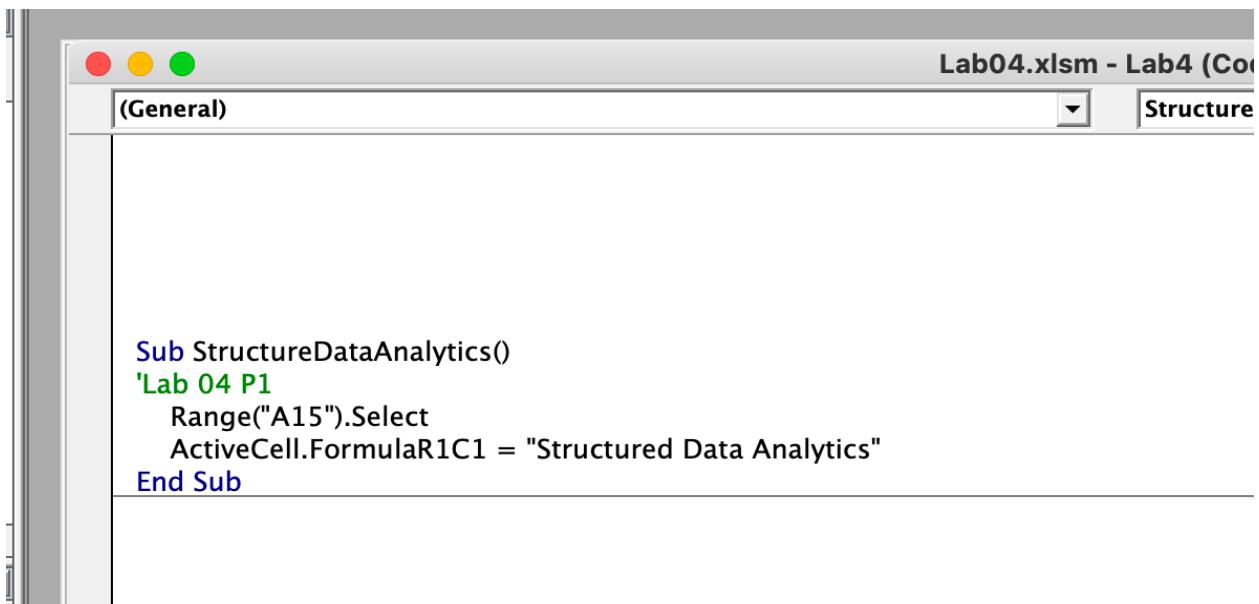
So I created the line chart using Excel.



Lab 04

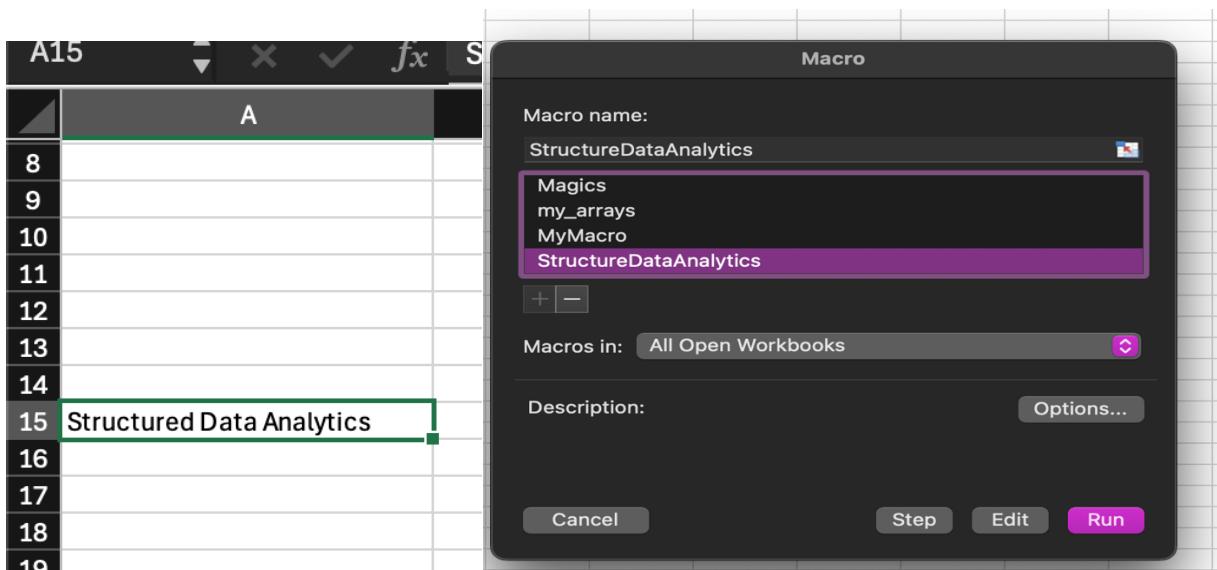
1. Record a macro of your choice. Submit screenshot of macro code and results

Recorded VBA:



```
Sub StructureDataAnalytics()
'Lab 04 P1
    Range("A15").Select
    ActiveCell.FormulaR1C1 = "Structured Data Analytics"
End Sub
```

Result:



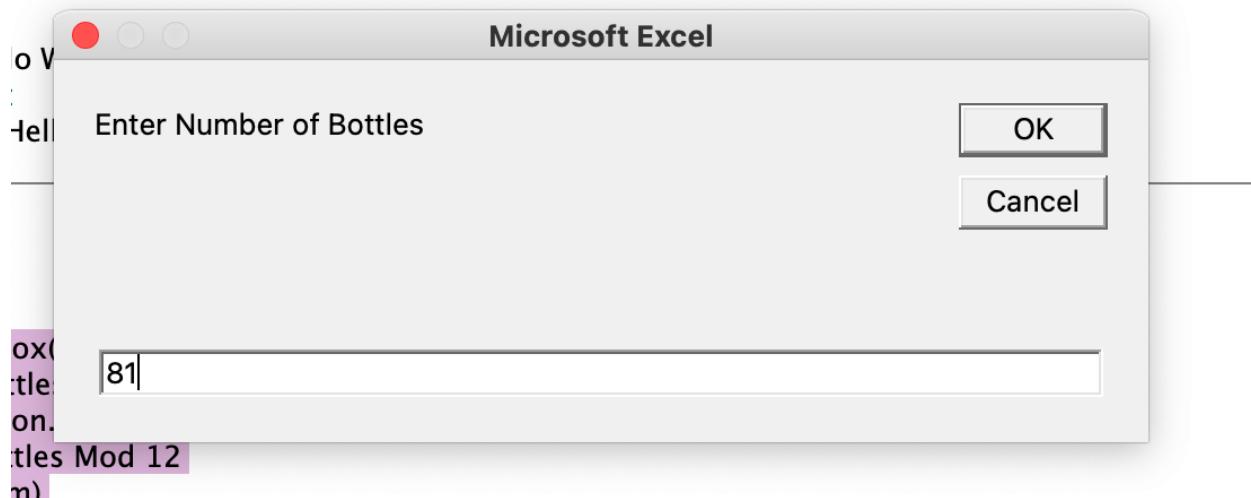
The screenshot shows a Microsoft Excel spreadsheet with a single cell selected, containing the text "Structured Data Analytics". The cell is highlighted with a green border. To the right of the spreadsheet, a "Macro" dialog box is open. The "Macro name:" field contains "StructureDataAnalytics". Below it, a list box shows several macro names, with "StructureDataAnalytics" highlighted in purple. The "Macros in:" dropdown is set to "All Open Workbooks". At the bottom of the dialog box are "Cancel", "Step", "Edit", and "Run" buttons.

2. Write a VBA sub for this problem: A small biz owner needs your help. The owner has 81 bottles that he must package into cases. Each case (box) can hold dozen bottles. How many empty cases (boxes) does the owner need, and how many bottles are left? Use a user input box for the owner to input the total number of bottles. Submit 3 screenshots: one showing code, one showing user input, one showing the result.

Code:

```
Sub boxes()
'Lab 4 P2
bottles = InputBox("Enter Number of Bottles")
Debug.Print (bottles)
num = Application.Quotient(bottles, 12)
remaining = bottles Mod 12
Debug.Print (num)
Debug.Print (remaining)
End Sub
```

User Input:



Result:

	81
	6
	9

3. Write a VBA sub that extracts the time from the string returned by the Now function and outputs it in a message box. Hint: search for VBA string functions

Code:

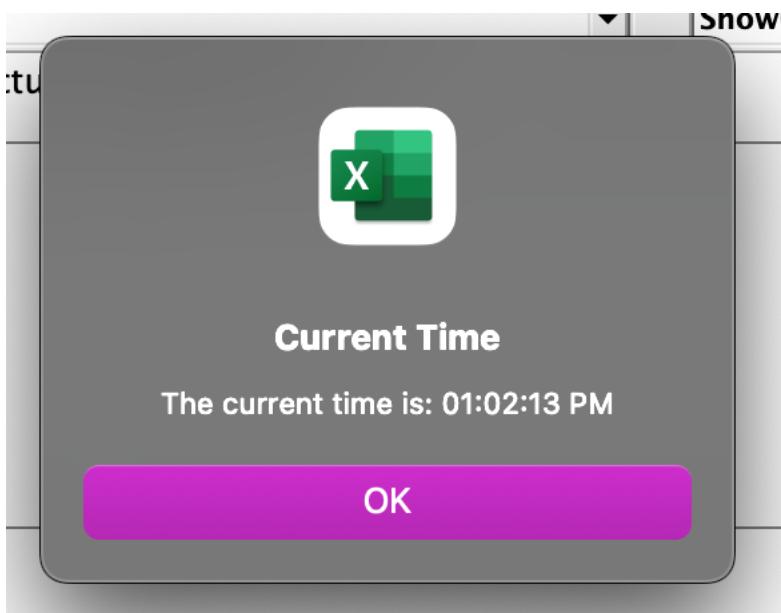
```
Sub ShowCurrentTime()
    Dim currentTime As String
    Dim currentDateTime As String

    ' Get the current date and time
    currentDateTime = Now

    ' Extract the time part from the date-time string
    currentTime = Format(currentDateTime, "hh:mm:ss AM/PM")

    ' Display the time in a message box
    MsgBox "The current time is: " & currentTime, vbInformation, "Current Time"
End Sub
```

Result:



Gen AI Prompt:

5. Gen AI Prompt Eng: ask you AI chatbot to solve this problem. Rub AI code. Compare AI code with your code. Which is better and why. Hint: it all in the prompt.

Prompt:

Write a VBA sub that automatically sums the rows, column and diagonal of 5X5 magic square. The sum is 65

Gen AI:

```
text
Sub CheckMagicSquare()
    Dim ws As Worksheet
    Dim rowSum(1 To 5) As Integer
    Dim colSum(1 To 5) As Integer
    Dim diag1Sum As Integer, diag2Sum As Integer
    Dim i As Integer, j As Integer

    Set ws = ActiveSheet

    ' Calculate row sums
    For i = 1 To 5
        rowSum(i) = 0
        For j = 2 To 6 ' Columns B to F
            rowSum(i) = rowSum(i) + ws.Cells(i + 1, j).Value
        Next j
        ws.Cells(i + 1, 7).Value = rowSum(i) ' Write sum in column G
    Next i

    ' Calculate column sums
    For j = 2 To 6 ' Columns B to F
        colSum(j - 1) = 0
        For i = 1 To 5
            colSum(j - 1) = colSum(j - 1) + ws.Cells(i + 1, j).Value
        Next i
        ws.Cells(7, j).Value = colSum(j - 1) ' Write sum in row 7
    Next j
```

```
' Calculate main diagonal sum (top-left to bottom-right)
diag1Sum = 0
For i = 1 To 5
    diag1Sum = diag1Sum + ws.Cells(i + 1, i + 1).Value
Next i
ws.Cells(7, 7).Value = diag1Sum

' Calculate secondary diagonal sum (top-right to bottom-left)
diag2Sum = 0
For i = 1 To 5
    diag2Sum = diag2Sum + ws.Cells(i + 1, 7 - i).Value
Next i
ws.Cells(7, 8).Value = diag2Sum

' Check if all sums equal 65
If Application.WorksheetFunction.CountIf(ws.Range("G2:H7"), 65) = 12
Then
    MsgBox "Valid magic square! All sums equal 65"
Else
    MsgBox "Not a valid magic square. Sums are not all equal to 65"
End If
End Sub
```

This VBA sub does the following:

1. It sums each row and places the result in column F (column 6).
2. It sums each column and places the result in row 6.
3. It calculates the sum of the main diagonal (top-left to bottom-right) and places the result in cell F6.
4. It calculates the sum of the secondary diagonal (top-right to bottom-left) and places the result in cell G6.
5. Finally, it checks if all sums are equal to 65, which is the expected sum for a 5x5 magic square [1](#) [2](#).

To use this sub:

1. Open your Excel workbook and press Alt + F11 to open the Visual Basic Editor.
2. Insert a new module and paste the code above into it.
3. Make sure your 5x5 magic square is in cells A1:E5 of the active sheet.
4. Run the sub by pressing F5 in the VBA editor or by calling it from a button or another macro.

The sub will calculate all sums and display them around the magic square. It will also show a message box indicating whether the square is a valid 5x5 magic square with sum 65.

Remember that a true 5x5 magic square should contain all numbers from 1 to 25 without repetition, and all rows, columns, and main diagonals should sum to 65 [1](#) [2](#).

Code:

```
Sub CheckMagicSquare()
    Dim ws As Worksheet
    Dim rowSum(1 To 5) As Integer
    Dim colSum(1 To 5) As Integer
    Dim diag1Sum As Integer, diag2Sum As Integer
    Dim i As Integer, j As Integer

    Set ws = ActiveSheet

    ' Calculate row sums
    For i = 1 To 5
        rowSum(i) = 0
        For j = 2 To 6 ' Columns B to F
            rowSum(i) = rowSum(i) + ws.Cells(i + 1, j).Value
        Next j
        ws.Cells(i + 1, 7).Value = rowSum(i) ' Write sum in column G
    Next i

    ' Calculate column sums
    For j = 2 To 6 ' Columns B to F
        colSum(j - 1) = 0
        For i = 1 To 5
            colSum(j - 1) = colSum(j - 1) + ws.Cells(i + 1, j).Value
        Next i
        ws.Cells(7, j).Value = colSum(j - 1) ' Write sum in row 7
    Next j

    ' Calculate main diagonal sum (top-left to bottom-right)
    diag1Sum = 0
    For i = 1 To 5
        diag1Sum = diag1Sum + ws.Cells(i + 1, i + 1).Value
    Next i

    ' Calculate secondary diagonal sum (top-right to bottom-left)
    diag2Sum = 0
    For i = 1 To 5
        diag2Sum = diag2Sum + ws.Cells(i + 1, 6 - i).Value
    Next i

    If rowSum(1) = 65 And rowSum(2) = 65 And rowSum(3) = 65 And rowSum(4) = 65 And rowSum(5) = 65 And colSum(1) = 65 And colSum(2) = 65 And colSum(3) = 65 And colSum(4) = 65 And colSum(5) = 65 And diag1Sum = 65 And diag2Sum = 65 Then
        MsgBox "The 5x5 magic square is valid with sum 65."
    End If
```

```

' Calculate main diagonal sum (top-left to bottom-right)
diag1Sum = 0
For i = 1 To 5
    diag1Sum = diag1Sum + ws.Cells(i + 1, i + 1).Value
Next i
ws.Cells(7, 7).Value = diag1Sum

' Calculate secondary diagonal sum (top-right to bottom-left)
diag2Sum = 0
For i = 1 To 5
    diag2Sum = diag2Sum + ws.Cells(i + 1, 7 - i).Value
Next i
ws.Cells(7, 8).Value = diag2Sum

' Check if all sums equal 65
If Application.WorksheetFunction.CountIf(ws.Range("G2:H7"), 65) = 12 Then
    MsgBox "Valid magic square! All sums equal 65"
Else
    MsgBox "Not a valid magic square. Sums are not all equal to 65"
End If
End Sub

```

Result:

A	B	C	D	E	F	G	H
	17	24	1	8	15		
	23	5	7	14	16	65	
	4	6	13	20	22	65	
	10	12	19	21	3	65	
	11	18	25	2	9	65	
						0	
	48	41	64	57	50	50	73
Gen AI: Question 5							

Comparison of Original and Gen-AI Code:

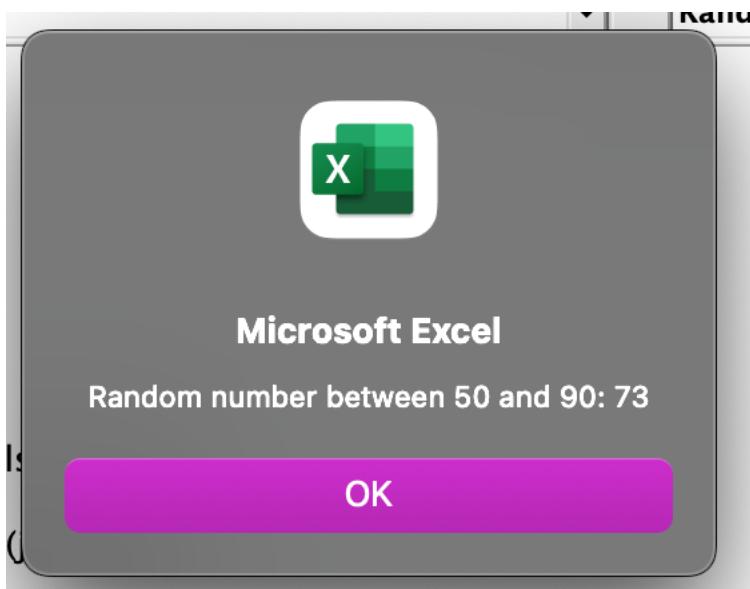
- Original code is more easy and gives the proper result.
- Where as the code given by Gen AI is very complex and hard to understand.
- The code given by Gen AI is not giving correct results which implies that AI models can make mistakes.
- On the other hand, code written in Problem 4 gave the correct result.

6. Write a VBA sub to return a random number between 50 and 90.

Code:

```
Sub RandomNumber50To90()
'Lab 4 P7
    Dim randomNumber As Integer
    randomNumber = Int((90 - 50 + 1) * Rnd + 50)
    MsgBox "Random number between 50 and 90: " & randomNumber
End Sub
```

Result:



7. Redo Problem 2, but write a VBA function. The input to the function is one cell value.
Submit 3 screenshots: one showing code, one showing Excel sheet input, one showing result.

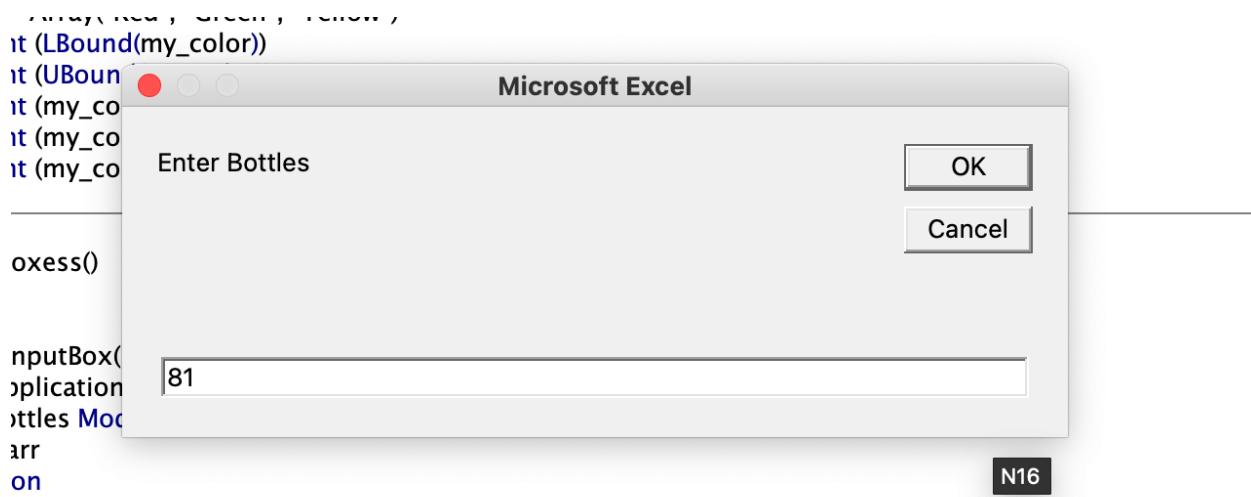
Code:

```
Function boxess()
'Lab 4 P7
    Dim arr(1)
    bottles = InputBox("Enter Bottles")
    arr(0) = Application.Quotient(bottles, 12)
    arr(1) = bottles Mod 12
    boxess = arr
End Function
```

Result:

	=boxess()	

User Input:



Result:

	6	9	

8. Create and use a custom object of your choice in Excel. Your custom object must have at least 3 attributes and 3 behaviors. Submit 2 screenshots: one showing code and showing results

Code:

Class Module:

```
(General)
' Class module: Car
Private pMake As String
Private pModel As String
Private pYear As Integer

' Property to set and get the Make
Public Property Get Make() As String
    Make = pMake
End Property

Public Property Let Make(Value As String)
    pMake = Value
End Property

' Property to set and get the Model
Public Property Get Model() As String
    Model = pModel
End Property

Public Property Let Model(Value As String)
    pModel = Value
End Property

' Property to set and get the Year
Public Property Get Year() As Integer
    Year = pYear
End Property

Public Property Let Year(Value As Integer)
    pYear = Value
End Property

' Method to display car details
Public Sub DisplayDetails()
    MsgBox "Make: " & pMake & vbCrLf & _
        "Model: " & pModel & vbCrLf & _
        "Year: " & pYear, vbInformation, "Car Details"

' Property to set and get the Year
Public Property Get Year() As Integer
    Year = pYear
End Property

Public Property Let Year(Value As Integer)
    pYear = Value
End Property

' Method to display car details
Public Sub DisplayDetails()
    MsgBox "Make: " & pMake & vbCrLf & _
        "Model: " & pModel & vbCrLf & _
        "Year: " & pYear, vbInformation, "Car Details"
End Sub

' Method to update the year
Public Sub UpdateYear(newYear As Integer)
    pYear = newYear
End Sub

' Method to check if the car is classic (older than 20 years)
Public Function IsClassic() As Boolean
    IsClassic = (Year(Date) - pYear) > 20
End Function
```

Subroutine:

```
Sub TestCar()
    Dim myCar As Car
    Set myCar = New Car

    ' Set attributes
    myCar.Make = "Toyota"
    myCar.Model = "Corolla"
    myCar.Year = 2010

    ' Call methods
    myCar.DisplayDetails

    ' Update year and check if classic
    myCar.UpdateYear 2000
    MsgBox "Is the car classic? " & IIf(myCar.IsClassic, "Yes", "No"), vbInformation, "Classic Check"
End Sub
```

Result:

