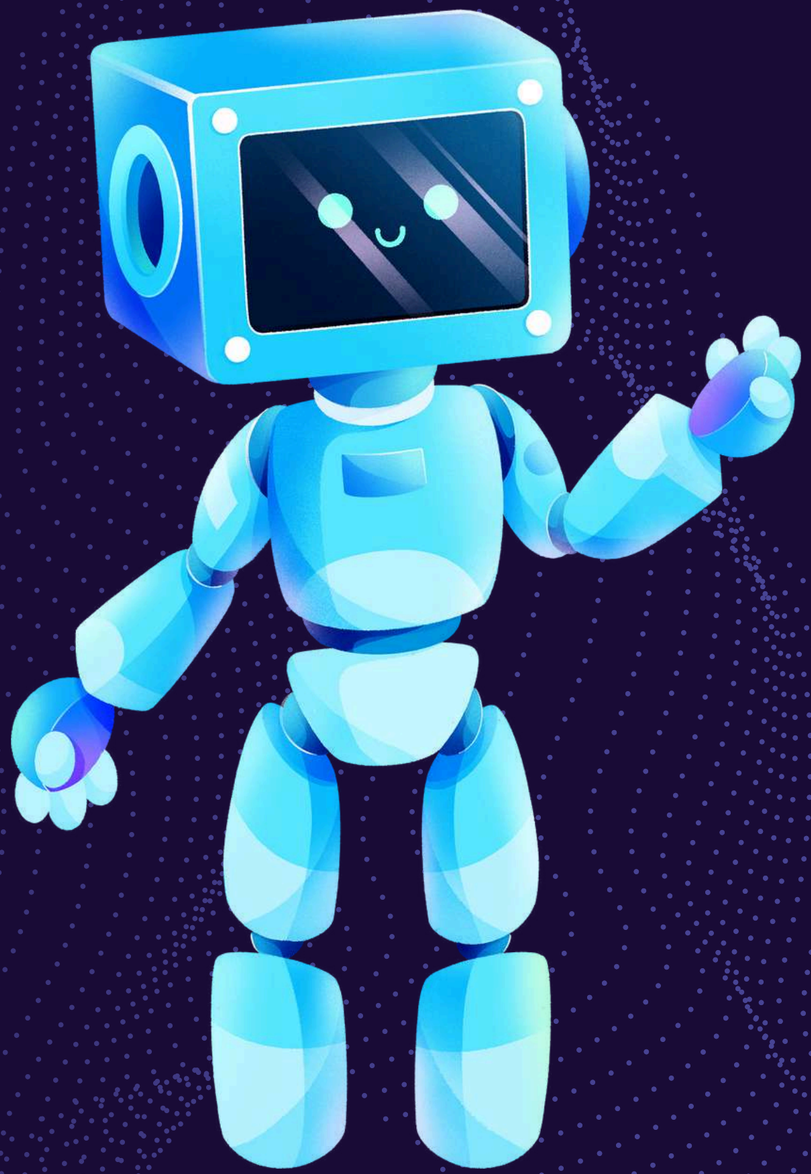


PROJECT 1 PRESENTATION

Forecasting Analysis Report

The project provides key financial and performance metrics from various companies, including company information, financial metrics, market indicators, risk and performance serving as a foundation for analysis, trend identification, and predictive modeling to support strategic decision-making.

Faaza Naima

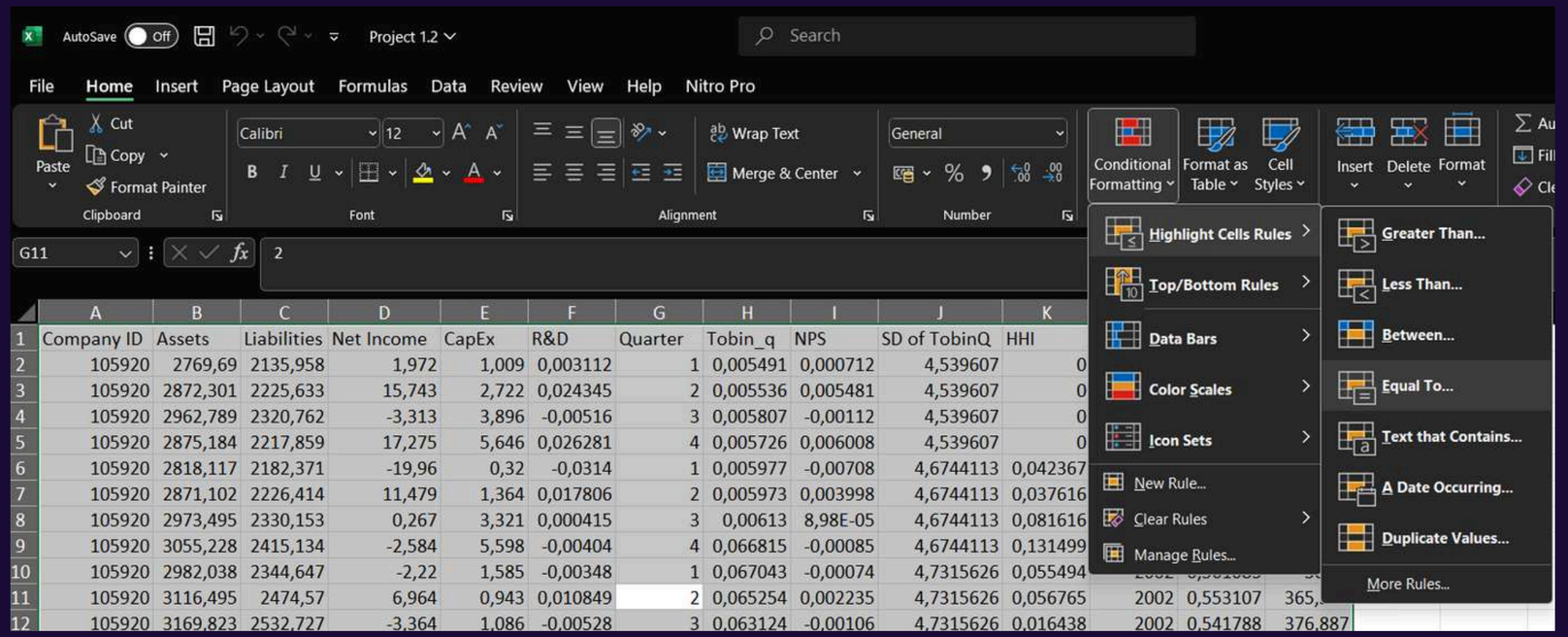
bit.ly/Project1-FTDA-Faaza-Naima

Overview of Financial Data Metrics

The financial dataset consists of various metrics from multiple companies, including:

Company ID : A unique identifier for each company
Assets : The total assets owned by the company
Liabilities : The total liabilities (debts) of the company
Net Income : The net income after all expenses have been deducted
CapEx : The capital spent on fixed asset investments
R&D : Expenses allocated for research and development activities
Quarter : The quarter within a year for which the data is recorded
Tobin_q : The market-to-book value ratio of the company
NPS : An indicator of customer loyalty towards the company
SD of TobinQ : The standard deviation of Tobin_q, measuring its volatility
HHI : A market concentration index used to measure industry competition
Year : The year associated with the financial data
AltmanZ : The Altman Z-score used to predict bankruptcy risk.
Sales : The total sales revenue of the company for a given period

TASK 1: Data Preparation & Cleaning Data



	A	B	C	D	E	F	G	H	I	J	K
1	Company ID	Assets	Liabilities	Net Income	CapEx	R&D	Quarter	Tobin_q	NPS	SD of TobinQ	HHI
2	105920	2769,69	2135,958	1,972	1,009	0,003112	1	0,005491	0,000712	4,539607	0
3	105920	2872,301	2225,633	15,743	2,722	0,024345	2	0,005536	0,005481	4,539607	0
4	105920	2962,789	2320,762	-3,313	3,896	-0,00516	3	0,005807	-0,00112	4,539607	0
5	105920	2875,184	2217,859	17,275	5,646	0,026281	4	0,005726	0,006008	4,539607	0
6	105920	2818,117	2182,371	-19,96	0,32	-0,0314	1	0,005977	-0,00708	4,6744113	0,042367
7	105920	2871,102	2226,414	11,479	1,364	0,017806	2	0,005973	0,003998	4,6744113	0,037616
8	105920	2973,495	2330,153	0,267	3,321	0,000415	3	0,00613	8,98E-05	4,6744113	0,081616
9	105920	3055,228	2415,134	-2,584	5,598	-0,00404	4	0,066815	-0,00085	4,6744113	0,131499
10	105920	2982,038	2344,647	-2,22	1,585	-0,00348	1	0,067043	-0,00074	4,7315626	0,055494
11	105920	3116,495	2474,57	6,964	0,943	0,010849	2	0,065254	0,002235	4,7315626	0,056765
12	105920	3169,823	2532,727	-3,364	1,086	-0,00528	3	0,063124	-0,00106	4,7315626	0,016438

- **Identifying and Handling Missing Values**

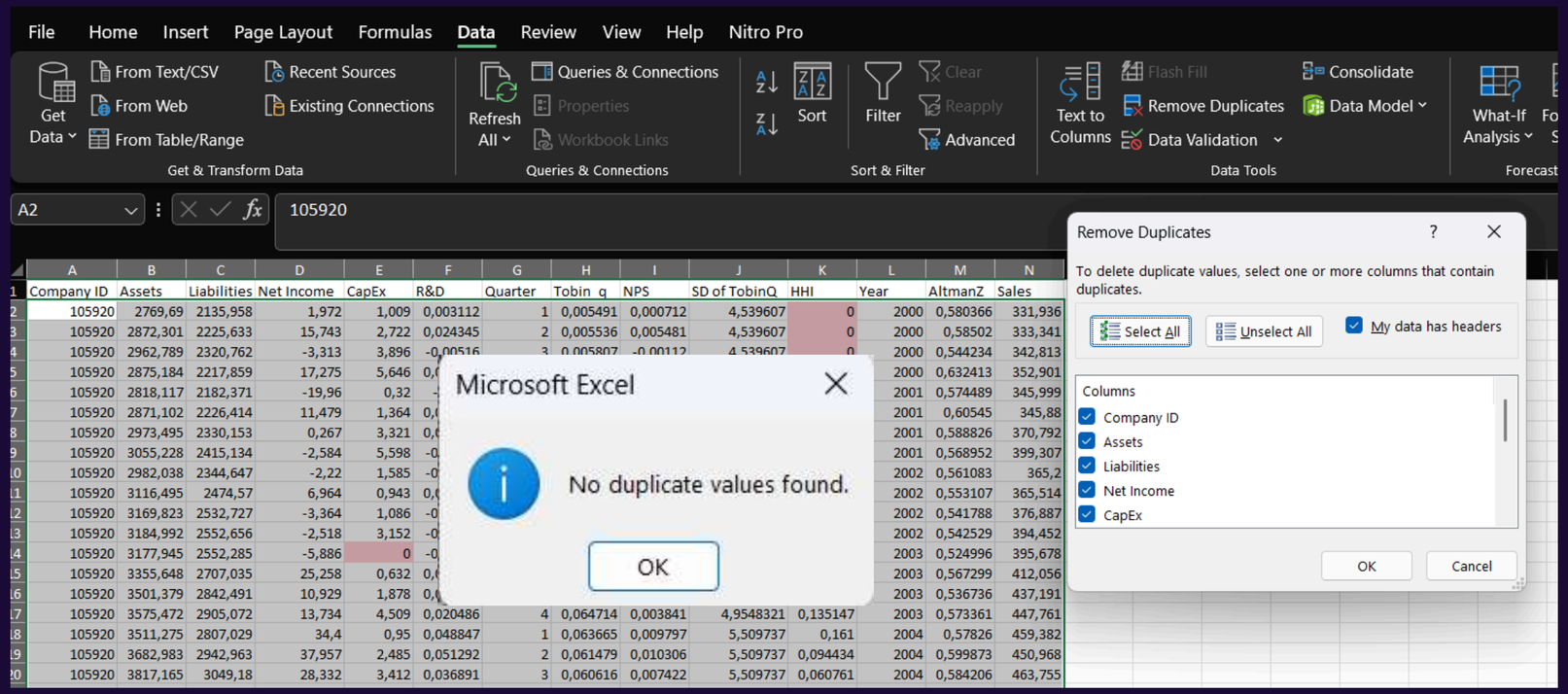
S2

=IF(I2=0; AVERAGEIF(A:A; A2; I:I); I2)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Company ID	Assets	Liabilities	Net Income	CapEx	R&D	Quarter	Tobin_q	Tobin_q	NPS	SD of TobinQ	HHI	HHI(old	Year	Altman2	Sales	Sales(c
2	1004	383,79	197,618	6,109	7,125	0,032814	1	2,8970468	2,8970468	0,015918	2,8584177	0,083666045	0	1990	1,605438	112,278	112,278
3	1004	388,521	198,973	6,224	18,415	0,032836	2	2,2059419	2,2059419	0,01602	2,8584177	0,083666045	0	1990	1,6241895	119,396	119,396
4	1004	385,752	191,734	6,697	1,734	0,034517	3	1,6374686	1,6374686	0,017361	3,9862406	0,083666045	0	1990	1,6752771	116,092	116,092
5	1004	382,978	193,475	0,126	3,348	0,000665	4	1,0769808	1,0769808	0,000329	3,9862406	0,083666045	0	1990	1,5944426	115,808	115,808
6	1004	383,354	191,593	3,977	5,621	0,020739	1	1,0020698	1,0020698	0,010374	3,9862406	0,049359627	0,04936	1991	1,6550452	117,82	117,82
7	1004	379,958	186,18	4,001	8,884	0,020647	2	1,1785454	1,1785454	0,01053	3,9862406	-0,021558512	-0,02156	1991	1,6825237	116,822	116,822
8	1004	376,618	181,542	3,444	1,301	0,017655	3	1,2686919	1,2686919	0,009145	4,1942592	-0,075397097	-0,0754	1991	1,6915777	107,339	107,339
9	1004	391,089	195,71	2,686	2,401	0,013748	4	1,1693237	1,1693237	0,006868	4,1942592	-0,11968085	-0,11968	1991	1,5907751	101,948	101,948
10	1004	396,971	199,895	2,88	5,984	0,014614	1	1,082652	1,082652	0,007255	4,1942592	-0,10948905	-0,10949	1992	1,5800571	104,92	104,92

V	W	X	Y
	1,6	21,0	0,5
Tidak Cocok	2,8970468	0,083666045	112,278
Tidak Cocok	2,2059419	0,083666045	119,396
Tidak Cocok	1,6374686	0,083666045	116,092
Tidak Cocok	1,0769808	0,083666045	115,808
Tidak Cocok	1,0020698	0,049359627	117,82
Tidak Cocok	1,1785454	-0,021558512	116,822
Tidak Cocok	1,2686919	-0,075397097	107,339
Tidak Cocok	1,1693237	-0,11968085	101,948

Handling Missing Values : Fill 0 values in a column with the average value for the same company (Company ID) for the columns HHI, Tobin_q, and sales.



Company ID	Assets	Liabilities	Net Income	CapEx	R&D	Quarter	Tobin_q	NPS	SD of TobinQ	HHI	Year	AltmanZ	Sales
105920	2769,69	2135,958	1,972	1,009	0,003112	1	0,005491	0,000712	4,539607	0	2000	0,580366	331,936
105920	2872,301	2225,633	15,743	2,722	0,024345	2	0,005536	0,005481	4,539607	0	2000	0,58502	333,341
105920	2962,789	2320,762	-3,313	3,896	-0,00516	3	0,005807	-0,00112	4,539607	0	2000	0,544234	342,813
105920	2875,184	2217,859	17,275	5,646	0,(						2000	0,632413	352,901
105920	2818,117	2182,371	-19,96	0,32	-(						2001	0,574489	345,999
105920	2871,102	2226,414	11,479	1,364	0,(						2001	0,60545	345,88
105920	2973,495	2330,153	0,267	3,321	0,(						2001	0,588826	370,792
105920	3055,228	2415,134	-2,584	5,598	-0						2001	0,568952	399,307
105920	2982,038	2344,647	-2,22	1,585	-0						2002	0,561083	365,2
105920	3116,495	2474,57	6,964	0,943	0,(						2002	0,553107	365,514
105920	3169,823	2532,727	-3,364	1,086	-0						2002	0,541788	376,887
105920	3184,992	2552,656	-2,518	3,152	-0						2002	0,542529	394,452
105920	3177,945	2552,285	-5,886	0	-0						2003	0,524996	395,678
105920	3355,648	2707,035	25,258	0,632	0,(						2003	0,567299	412,056
105920	3501,379	2842,491	10,929	1,878	0,(						2003	0,536736	437,191
105920	3575,472	2905,072	13,734	4,509	0,020486	4	0,064714	0,003841	4,9548321	0,135147	2003	0,573361	447,761
105920	3511,275	2807,029	34,4	0,95	0,048847	1	0,063665	0,009797	5,509737	0,161	2004	0,57826	459,382
105920	3682,983	2942,963	37,957	2,485	0,051292	2	0,061479	0,010306	5,509737	0,094434	2004	0,599873	450,968
105920	3817,165	3049,18	28,332	3,412	0,036891	3	0,060616	0,007422	5,509737	0,060761	2004	0,584206	463,755

Checking duplicate

Cek Outlier	
IQR	346,834
Q1	45,2283
Q2	119,116
Q3	392,062
Lower Bond	-475,023
Upper Bond	912,313

1. Checking outlier in Sales Data

```
=IF(AND(P2>0;P2<912314); "";"OUTLIER")
```

2. Tagging outlier

```
=COUNTIF(S2:S53076; "OUTLIER")
```

3. Counting outlier

```
=COUNTIF(S2:S53076; "OUTLIER")
```

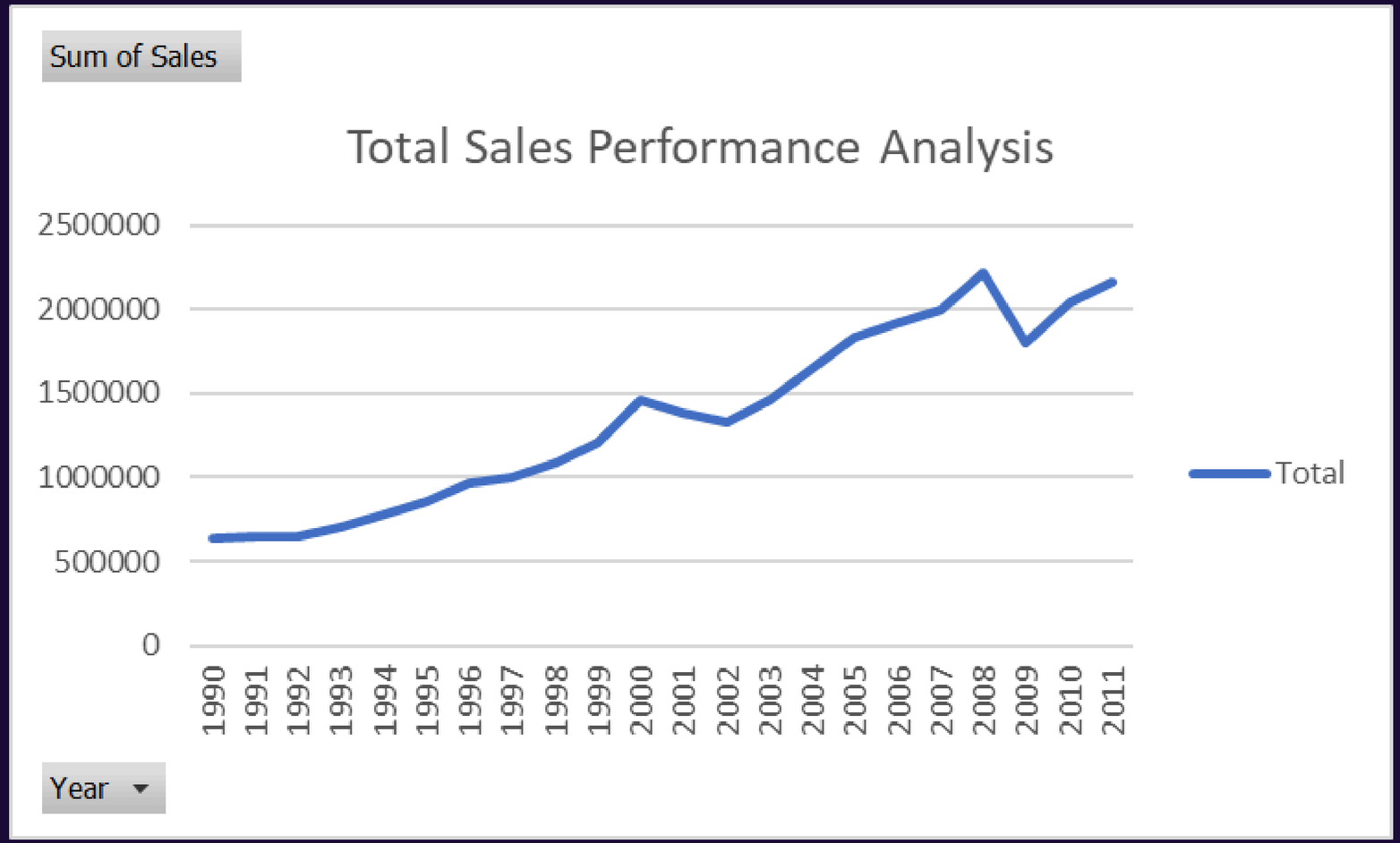
OUTLIER	0

Outliers Analysis in Sales Data

- Handling missing values for companies that have no sales value at all across all years. Additionally, 98% of the historical data lacks z-score values. I applied a pivot table to identify companies with a total sales value of 0, then used the formula =IF(ISNUMBER(MATCH(A2, sales0!$A$1:$A$1966, 0)), "Match", "No Match"). If it did not match, the entry was removed. As a result of this process, the dataset was reduced from 391,470 to 53,067 entries.
- Handling missing value: Fill 0 values in a column with the average value for the same company (Company ID) for the columns HHI, Tobin_q, and sales. "If the result remains 0 due to the absence of data for the company, the default value for Tobin's Q will be set to 1. This is a proxy approximation for Tobin's Q, as outlined in Financial Management, Vol. 23, No. 3, Autumn 1994, pages 70-74. Page 2. CHUNG & PRUITT / A SIMPLE APPROXIMATION OF TOBIN'S Q 71. Meanwhile, for HHI, the default value will be set to 0.5, based on the industry average.

Handling Activities

TASK 2: Overall Description of Sales Data



General Overview and Insights from Sales Data

- The line graph illustrates the total sales performance from 1991 to 2011. Sales experienced a general upward trend throughout the analyzed period, with some fluctuations observed. Notably, there was a significant increase in sales from 2006 to 2008, followed by a slight decline in 2009. However, sales rebounded and reached a peak in 2011.
- This growth can be attributed to several factors, such as new product launches, increased market penetration, and economic growth increasing from 5.5% to 6.1%. The temporary dip in sales in 2009 aligns with the global financial crisis of 2008, which caused economic growth to decline from 6.01% to 4.63%. However, the subsequent recovery highlights the company's resilience and ability to adapt to changing market conditions.

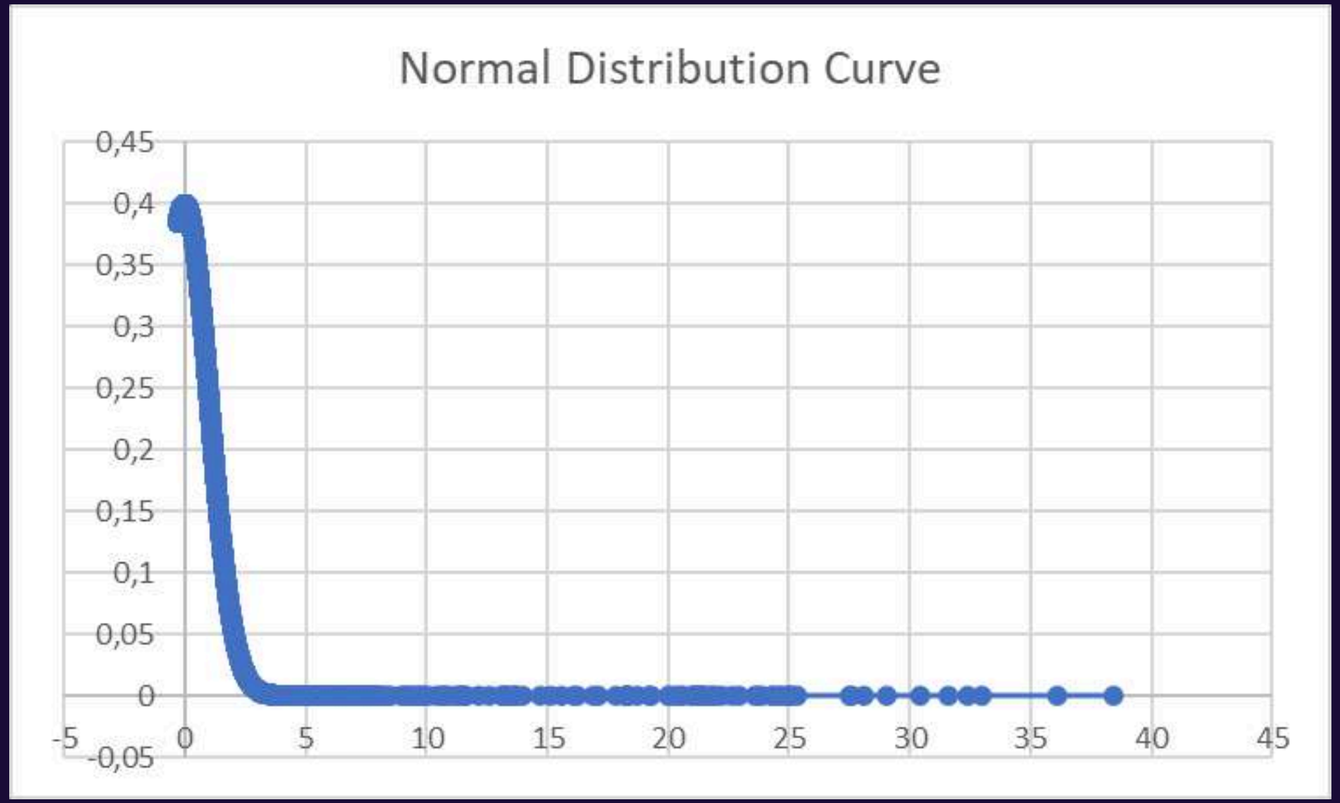
source: https://databoks.katadata.co.id/ekonomi-makro/statistik/7ea9b6fed5d8362/inilah-pertumbuhan-ekonomi-indonesia-sejak-1961

Insight

Sales	
Mean	561,6763
Standard Error	8,778786
Median	119,055
Mode	10,13654
Standard Deviation	2022,458
Sample Variance	4090335
Kurtosis	397,0395
Skewness	16,559
Range	78310
Minimum	0,001
Maximum	78310
Sum	29810970
Count	53075

- The sales data has a mean of 561.68, with a standard error of 8.78, indicating the precision of the mean estimate.
- The median sales value is 119.06, suggesting that half of the data is below this value and half is above it. The mode, or most frequently occurring sales value, is 10.14.
- The high standard deviation of 2022.46 reflects a large spread of data, while the sample variance of 4,090,335.36 indicates significant dispersion.
- The kurtosis value of 397.04 is very high, indicating a sharp distribution with many outliers, while the skewness of 16.56 suggests a strong rightward skew with many small values and fewer large ones.
- The range of the data is 78,310, with a minimum value of 0.001 and a maximum of 78,310. The total sales amount is 29,810,970.25, spread across 53,075 data entries.

Data Analysis: Statistic Descriptive



Identifying Sales Data Distribution: Skewed or Normal?

- The curve is **skewed to the right**, indicating that the distribution is not perfectly normal.
- The peak of the curve appears to be around 0, suggesting that the mean of the distribution is close to 0.
- The curve is relatively wide, suggesting a relatively high standard deviation.

Key Takeaways

TASK 3: Developing Hypotheses Based on Features

Regression Statistics	
Multiple R	0,84494
R Square	0,71393
Adjusted R Square	0,71387
Standard Error	1081,84
Observations	53075

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	1,5E+11	1,3E+10	11035,3	0
Residual	53062	6,2E+10	1170388		
Total	53074	2,2E+11			

	Coefficients	*standard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	8322,68	1942,12	4,28536	1,8E-05	4516,10676	12129,25713	4516,10676	12129,25713
Assets	0,37547	0,0024	156,52	0	0,370767096	0,380170662	0,370767096	0,380170662
Liabilities	-0,36919	0,0026	-141,843	0	-0,374294913	-0,364091739	-0,374294913	-0,364091739
Net Income	0,90078	0,02197	41,005	0	0,857719039	0,943831912	0,857719039	0,943831912
CapEx	0,5281	0,01432	36,8777	6E-294	0,500036271	0,556172537	0,500036271	0,556172537
R&D	-1,08379	32,1862	-0,03367	0,97314	-64,16895238	62,00137865	-64,16895238	62,00137865
Quarter	-10,5787	4,22142	-2,50596	0,01221	-18,85270879	-2,304684573	-18,85270879	-2,304684573
Tobin_q	0,82118	0,99398	0,82616	0,40872	-1,127029953	2,769396948	-1,127029953	2,769396948
NPS	1623,29	221,441	7,33061	2,3E-13	1189,269297	2057,320136	1189,269297	2057,320136
SD of TobinQ	-0,56651	12,9267	-0,04382	0,96504	-25,90287722	24,76986332	-25,90287722	24,76986332
HHI	17,3767	14,6466	1,1864	0,23547	-11,33072815	46,0841924	-11,33072815	46,0841924
Year	-4,11933	0,9893	-4,16386	3,1E-05	-6,05837405	-2,180283176	-6,05837405	-2,180283176
AltmanZ	-0,02828	0,03858	-0,7332	0,46344	-0,103896047	0,047326693	-0,103896047	0,047326693

Data Analysis: Hypotheses Based on Features

- The p-value helps determine the statistical significance of each coefficient. It tests the null hypothesis that the coefficient is equal to 0 (i.e., the independent variable has no effect on the dependent variable).
- P-value < 0.05: This suggests that the coefficient is statistically significant (reject the null hypothesis).
- P-value ≥ 0.05: This suggests that the coefficient is not statistically significant (fail to reject the null hypothesis).
- Variables such as Assets, Liabilities, Net Income, CapEx, Quarter, NPS, and Year have statistically significant relationships with Sales, as their p-values are less than 0.05.
- Variables like R&D, Tobin_q, SD of TobinQ, HHI, and AltmanZ are not statistically significant, meaning they do not have a meaningful effect on Sales based on this regression model.

Key Takeaways

Column1	Assets	Liabilities	Net Income	CapEx	R&D	Quarter	Tobin_q	NPS	SD of Tobin	HHI	Year	AltmanZ	Sales
Assets	1												
Liabilities	0,973147	1											
Net Income	0,300833	0,19494127	1										
CapEx	0,453662	0,30345113	0,324318372	1									
R&D	0,017085	0,01337536	0,099737764	0,007875	1								
Quarter	0,001411	0,00076831	-0,011173544	0,0774	-0,01272	1							
Tobin_q	0,02518	0,02108465	0,021999106	0,010595	0,028677	0,005446	1						
NPS	0,003025	-0,0108701	0,132519429	0,03207	0,306241	-0,008106	0,045012	1					
SD of Tobin	0,045745	0,03672566	0,039975025	0,046534	0,005626	0,007066	0,017617	0,029372	1				
HHI	-0,0084	-0,0067149	0,016630132	-0,00277	0,040498	0,006865	-0,007631	0,09016	0,065280924	1			
Year	0,071186	0,05245883	0,045439463	0,094735	-0,01645	-0,006625	-0,016551	0,006412	0,613228086	0,008553	1		
AltmanZ	-0,0053	-0,004885	-0,001297057	-0,00382	0,001649	0,004868	0,010069	0,013654	-0,01031521	0,000761	-0,01671	1	
Sales	0,562795	0,4055172	0,507801445	0,67455	0,0332	0,004976	0,025866	0,068433	0,041785038	-0,00131	0,074666	-0,004979	1

The correlation matrix provides insights into the relationships between various financial variables. For instance, there is a strong positive correlation between sales and net income, suggesting that an increase in sales is likely to lead to an increase in profits. Sales shows moderate positive correlations with assets, CapEx, and net income, but weak correlations with other variables such as R&D, NPS, Tobin's Q, and Altman Z.

Relationships Between Features and Sales: Positive, Negative, or No Correlation

TASK 4: Sales Forecasting Using Regression Analysis

Regression Statistics	
Multiple R	0,84494
R Square	0,71393
Adjusted R Square	0,71387
Standard Error	1081,84
Observations	53075

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	12	1,5E+11	1,3E+10	11035,3	0
Residual	53062	6,2E+10	1170388		
Total	53074	2,2E+11			

	Coefficients	*standard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	8322,68	1942,12	4,28536	1,8E-05	4516,10676	12129,25713	4516,10676	12129,25713
Assets	0,37547	0,0024	156,52	0	0,370767096	0,380170662	0,370767096	0,380170662
Liabilities	-0,36919	0,0026	-141,843	0	-0,374294913	-0,364091739	-0,374294913	-0,364091739
Net Income	0,90078	0,02197	41,005	0	0,857719039	0,943831912	0,857719039	0,943831912
CapEx	0,5281	0,01432	36,8777	6E-294	0,500036271	0,556172537	0,500036271	0,556172537
R&D	-1,08379	32,1862	-0,03367	0,97314	-64,16895238	62,00137865	-64,16895238	62,00137865
Quarter	-10,5787	4,22142	-2,50596	0,01221	-18,85270879	-2,304684573	-18,85270879	-2,304684573
Tobin_q	0,82118	0,99398	0,82616	0,40872	-1,127029953	2,769396948	-1,127029953	2,769396948
NPS	1623,29	221,441	7,33061	2,3E-13	1189,269297	2057,320136	1189,269297	2057,320136
SD of TobinQ	-0,56651	12,9267	-0,04382	0,96504	-25,90287722	24,76986332	-25,90287722	24,76986332
HHI	17,3767	14,6466	1,1864	0,23547	-11,33072815	46,0841924	-11,33072815	46,0841924
Year	-4,11933	0,9893	-4,16386	3,1E-05	-6,05837405	-2,180283176	-6,05837405	-2,180283176
AltmanZ	-0,02828	0,03858	-0,7332	0,46344	-0,103896047	0,047326693	-0,103896047	0,047326693

Data Analysis: Regression

- Given an R^2 value of 0.7139, it means that approximately 71.39% of the variance in Sales can be explained by the independent variables included in the model (such as Assets, Liabilities, Net Income, CapEx, R&D, etc.).
- F = 11035.32331: The F-statistic is very large, suggesting that the independent variables in the model significantly explain the variation in the dependent variable.
- Significance F = 0: The p-value of 0 indicates that the regression model is statistically significant. It means that there is strong evidence to reject the null hypothesis that the regression coefficients are equal to zero (no effect). A p-value of 0 suggests a very strong relationship between the independent variables and the dependent variable.

Key Takeaways

TASK 5: Dashboard

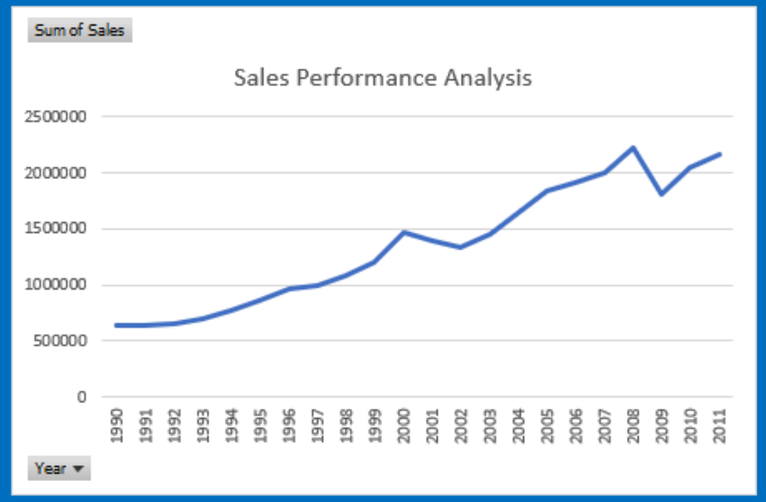
Forecasting Analysis Report

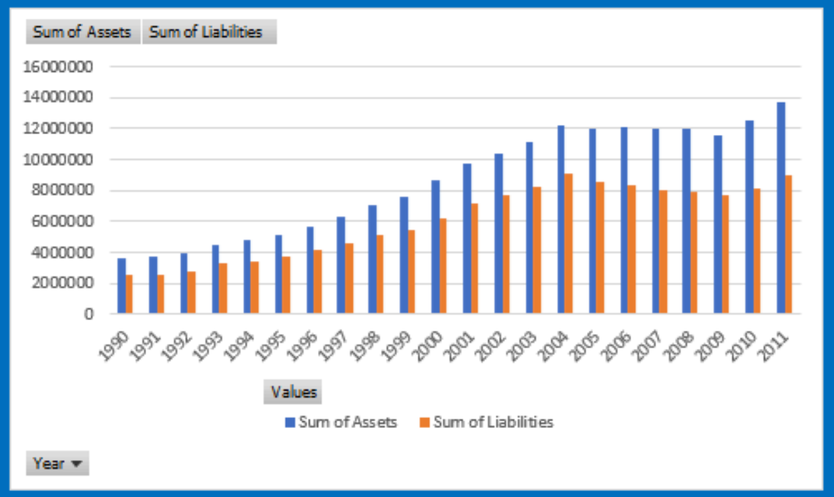
Company ID
- 2965
- 2967
- 2974
- 2975

Year
- 1990
- 1991
- 1992
- 1993

Quarter
- 1
- 2
- 3
- 4

ASSETS	LIABILITIES	NET INCOME	SALES
190505766,9	133928675,2	1436477,553	29781312,85





Dashboard

Terima Kasih!