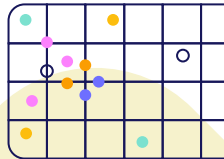
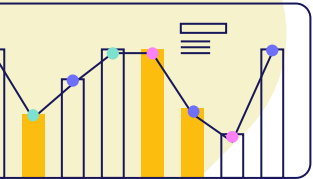


# ***Assignment 4 Forecasting Analysis Report***

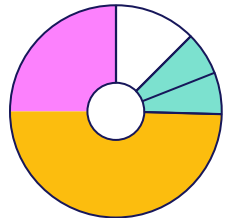
***Indah Ayu Permatasari***



# 01. *Business Understanding*

*Pada bagian ini, penulis hanya berfokus pada poin Business Task. Secara umum, Business Task yang ingin diselesaikan yaitu diantaranya :*

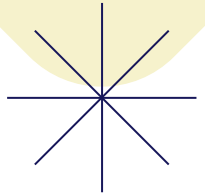
- ☐ *Memperoleh insight mengenai kondisi sales secara umum*
- ☐ *Menemukan insight dari data agar bisa meningkatkan performance dari sales*
- ☐ *Menemukan pengaruh berbagai fitur terhadap kondisi sales*



○

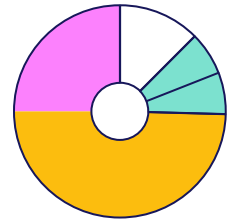
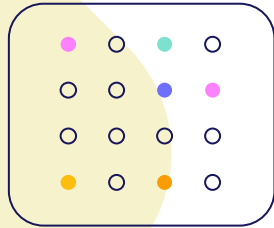
**02.**

# ***Data Understanding***



## ***2. 1. Data Collection***

*Data telah disediakan. Untuk mengakses data silakan klik [Link Data](#).*



○

## 2. 2 Data Description

Data terdiri dari 14 kolom dan 391.470 baris (termasuk label data).

Detail kolom :

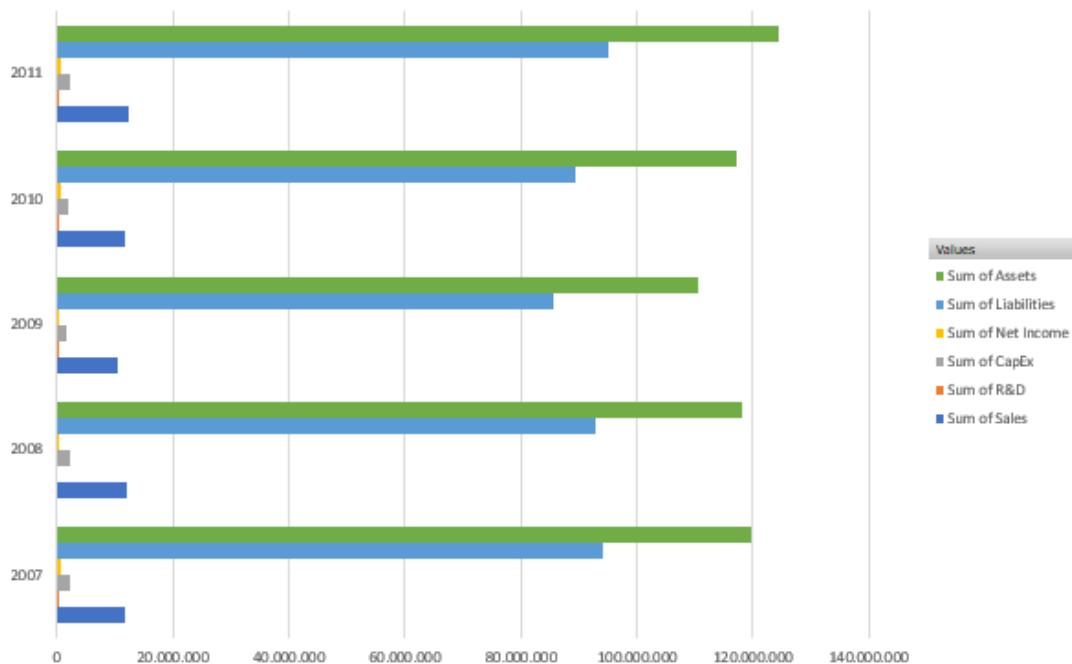
- Company ID : berisi ID dari Perusahaan (type : text)
- Assets : Jumlah asset yang dimiliki Perusahaan dimana asset merupakan sumber-sumber daya bernilai ekonomi milik pribadi/perusahaan dan diharapkan bisa menghasilkan keuntungan di masa mendatang. (type : currency , unit : dollar)
- Liabilities : Jumlah liabilitas yang dimiliki Perusahaan dimana liabilitas itu sendiri adalah kewajiban yang harus dibayar oleh sebuah perusahaan pada pihak yang bersangkutan dengan cara mengeluarkan sejumlah dana atau sumber daya ekonomi perusahaan tersebut. (type : currency , unit : dollar)
- Net Income : Jumlah laba bersih yang diperoleh Perusahaan (jumlah pendapatan yang tersisa setelah dikurangi semua biaya.) (type : currency , unit : dollar)
- CapEx : Dana yang digunakan oleh perusahaan untuk memperoleh, meningkatkan, dan memelihara aset fisik seperti properti, pabrik, bangunan, teknologi, atau peralatan. (type : currency , unit : dollar)
- R&D : R&D expenses (type : currency , unit : dollar)
- Quarter : berisi nilai quarter 1,2,3,atau 4 (type : number)
- Tobin\_q : Tobin's Q merupakan salah satu ratio dalam mengukur nilai perusahaan, rasio dari nilai pasar dari aset perusahaan (nilai pasar dari saham beredar ditambah hutang) dibagi dengan biaya penggantian aset perusahaan atau nilai buku. (type : number)
- NPS : Net promoter score (NPS) adalah skor yang hasilnya mencerminkan loyalitas konsumen. (type : number)
- SD of TobinQ : SD dari TobinQ (type : number)
- HHI : Indeks Herfindahl-Hirschman (HHI) adalah ukuran umum konsentrasi pasar dan digunakan untuk menentukan daya saing pasar (type : number)
- Year : berisi tahun recording data tiap rows (type : number)
- AltmanZ : **Altman Z-score** adalah suatu alat yang digunakan untuk memprediksi tingkat kebangkrutan suatu Perusahaan
- Sales : Jumlah sales (type : currency , unit : dollar)

## 2. 3. Data Exploration

	Total (\$)	Average (\$)	Max (\$)	Min (\$)
Assets	1.808.922.282	4.621	1.313.867	100
Liabilities	1.428.891.760	3.650	1.210.555	0
Net Income	8.980.669	23	26.615	-66.048
CapEx	31.342.824	80	27.239	0
R&D	6.112	0	2	-2
Sales	181.604.595	464	132.872	0

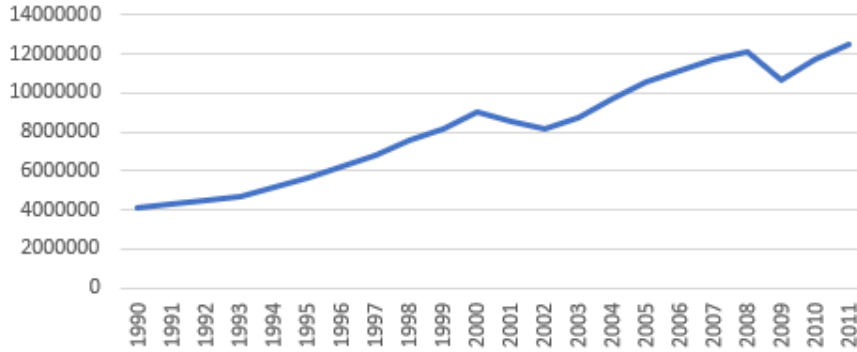
Tahun	Sum of Sales	Sum of R&D	Sum of CapEx	Sum of Net Income	Sum of Liabilities	Sum of Assets
2007	11.720.342	305	2.182.864	706.742	94.276.373	119.604.637
2008	12.141.158	-54	2.308.492	257.921	92.985.987	118.227.162
2009	10.654.019	92	1.766.369	462.886	85.645.798	110.572.572
2010	11.693.537	256	1.848.380	733.796	89.468.288	117.159.483
2011	12.529.705	303	2.168.929	804.617	95.159.155	124.478.539
<b>Grand Total</b>	<b>58.738.761</b>	<b>902</b>	<b>10.275.034</b>	<b>2.965.963</b>	<b>457.535.600</b>	<b>590.042.393</b>

Total Asset, Liabilities, Net Income, CapEx, R&D , Sales di 5 Tahun terakhir

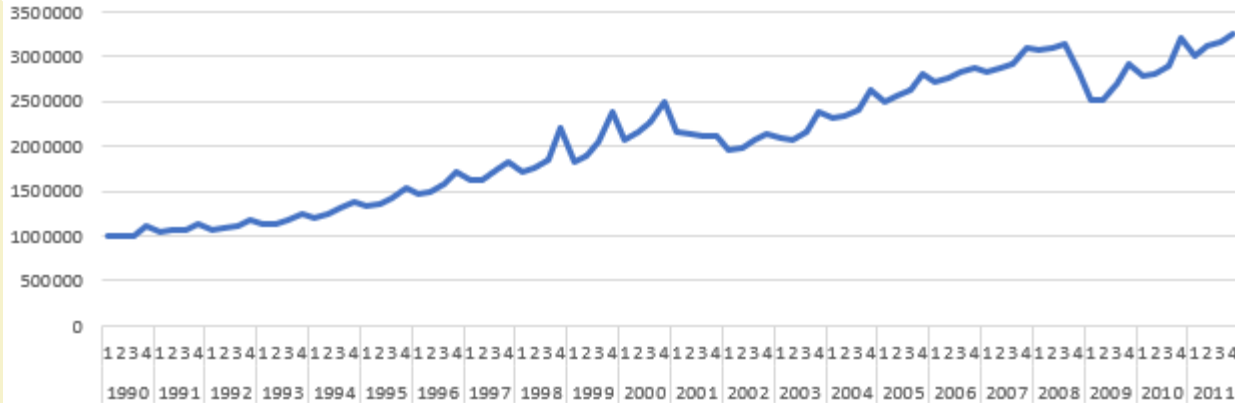


Dapat dilihat bahwa secara umum nilai jumlah asset merupakan nilai yang paling tinggi serta liabilitiesnya juga cenderung tinggi. Sementara itu, untuk net income sangat kecil serta nilai jumlah sales juga cenderung kecil. Biaya untuk RnD dan CapEx sangat kecil.

**Total Sales**



**Total Sales Quarterly**



Dapat dilihat bahwa secara umum terdapat peningkatan jumlah sales dari tahun 1990 hingga 2011. Ketika dilihat secara periode quarter nilainya naik turun tetapi masih cenderung menaik.

### *Top 5 Company*

Company ID	Jumlah Transaksi
2751	90
10801	89
5639	89
1602	89
29555	89

Company ID	Jumlah Sales (\$)
11259	4.699.290
2991	2.249.403
8549	1.539.787
5606	1.356.464
2136	1.238.199



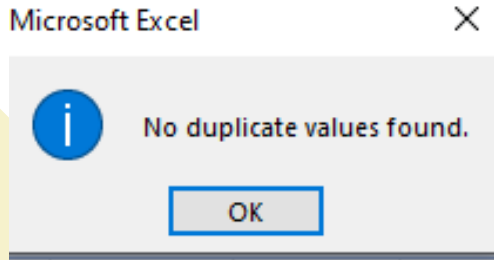
# 03. Data Preparation

## - Cek Missing Value

Tidak terdapat missing value (asumsi missing value disini hanya jika terdapat empty cell dan atau nilai n/a)

## - Cek Duplikasi Data

Tidak terdapat duplikasi data



## - Cek Outlier

Outlier Calculation	
Median	89.106
1st Quartile	23.613
3rd Quartile	301.8
IQR	278.187
Fence Multiplier	1.5
Lower	-393.668
Upper	719.0805

Terdapat 50.129 data outlier

## - Cek Korelasi Data

	Assets	Liabilities	Net Income	CapEx	R&D	Tobin_q	NPS	D of TobinQ	HHI	AltmanZ	Sales
Assets	1										
Liabilities	0.998847	1									
Net Income	0.414455	0.410041	1								
CapEx	0.070973	0.059085	0.01874	1							
R&D	0.008081	0.008018	0.129367	-0.00534	1						
Tobin_q	0.026259	0.024667	0.028621	0.029691	0.016108	1					
NPS	-0.01656	-0.01971	0.149441	0.008991	0.30297	0.039655	1				
SD of Tobin	0.012297	0.009498	0.02415	0.026381	0.003612	0.05096	0.015077	1			
HHI	-0.02045	-0.02269	0.016162	0.052997	0.045252	0.014247	0.134983	0.068963	1		
AltmanZ	-0.00367	-0.00425	0.000996	-0.00574	0.00178	0.002739	0.013109	-0.00428	-0.00347	1	
Sales	0.013223	-0.00367	0.062496	0.272762	0.053108	-0.00754	0.147949	0.006793	0.06395	-0.01452	1

Terlihat pada table bahwa korelasi fitur yang ada terhadap sales cenderung kecil (paling tinggi hanya 0.2 yaitu pada nilai CapEx. Untuk NPS hanya memiliki korelasi sebesar 0.14 serta fitur lainnya dibawah 0.1.

## ○ - *Deskriptive Statistics*

Data diambil setelah dikurangi dengan data outlier.

## - *Distribusi Data*

*Positively Skewed (skewed right)*  
*karena nilai mean > median*

Sales	
Mean	130.4535
Standard Error	0.273631
Median	66.785
Mode	0
Standard Deviation	159.8668
Sample Variance	25557.41
Kurtosis	2.178903
Skewness	1.662074
Range	719.056
Minimum	0
Maximum	719.056
Sum	44528982
Count	341340

# 04.

# Modeling & Evaluation

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	730.4659964	105.8885618	6.898441	5.27E-12	522.9274929	938.0045	522.927493	938.0044999
Assets	0.065453513	0.000361758	180.9316	0	0.064744477	0.066163	0.06474448	0.066162549
Liabilities	-0.072173324	0.000396952	-181.819	0	-0.07295134	-0.0714	-0.07295134	-0.07139531
Net Income	0.031060322	0.003052604	10.17502	2.58E-24	0.025077307	0.037043	0.02507731	0.037043338
CapEx	0.328153935	0.002635515	124.5122	0	0.322988401	0.333319	0.3229884	0.333319469
R&D	17.01165457	1.853855278	9.176366	4.48E-20	13.37815211	20.64516	13.3781521	20.64515703
Q1	6.918726656	0.703289486	9.837665	7.8E-23	5.540299705	8.297154	5.5402997	8.297153607
Q2	3.454886813	0.700234171	4.933902	8.06E-07	2.082448191	4.827325	2.08244819	4.827325435
Q3	1.018588551	0.69990638	1.455321	0.145581	-0.35320761	2.390385	-0.35320761	2.390384713
Tobin_q	-1.007914122	0.056159912	-17.9472	5.45E-72	-1.11798592	-0.89784	-1.11798592	-0.89784233
NPS	877.6212125	12.40637936	70.73951	0	853.3050695	901.9374	853.305069	901.9373554
SD of TobinQ	-3.80166423	0.693544174	-5.4815	4.22E-08	-5.16099065	-2.44234	-5.16099065	-2.44233781
HHI	12.80573415	0.848924991	15.08465	2.12E-51	11.14186584	14.4696	11.1418658	14.46960246
Year	-0.307879574	0.053959613	-5.70574	1.16E-08	-0.41363885	-0.20212	-0.41363885	-0.2021203
AltmanZ	-0.034542963	0.002917663	-11.8393	2.48E-32	-0.0402615	-0.02882	-0.0402615	-0.02882443

Hasil dari regresi ditemukan bahwa hanya fitur Intercept, Net Income, R&D, Q1, Q2, Tobin\_q, SD of TobinQ, HHI, year, dan Altman yang memiliki p-value < 0.05 yang artinya statistical signifikan.

R square yang diperoleh sangat kecil yaitu 0.17 ini mengindikasikan bahwa hanya terdapat 17% keragaman data yang dapat dijelaskan oleh fitur-fitur yang berpengaruh signifikan.

Regression Statistics	
Multiple R	0.423480472
R Square	0.17933571
Adjusted R Square	0.179302049
Standard Error	144.8271836
Observations	341340