



01

FORBES

Informasi tentang
perusahaan besar di dunia

Pemrosesan Awal

Import packages

```
library(dplyr)
```

Untuk memanipulasi *data frame*

```
library(DataExplorer)
```

Untuk membuat *plot* dari persentase *missing value*

```
library(ggplot2)
```

Untuk membuat visualisasi data

```
library(ggrepel)
```

Untuk membuat *label* pada visualisasi dinamis

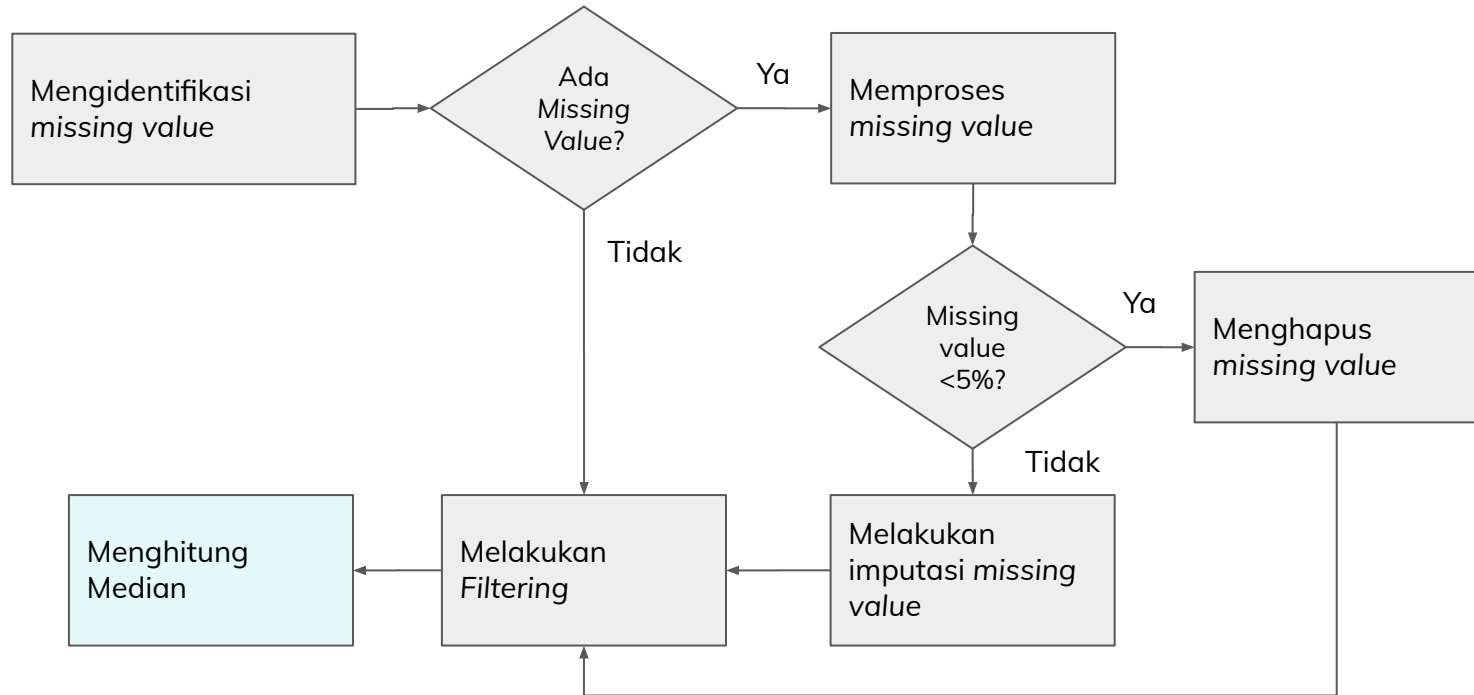
Import Data

```
forbes <-  
read.csv(file.choose())
```

Membuat variabel 'forbes' sebagai nama dataset yang akan digunakan. Data dipilih manual setelah ada tampilan *window* baru.

Nomor 1.1

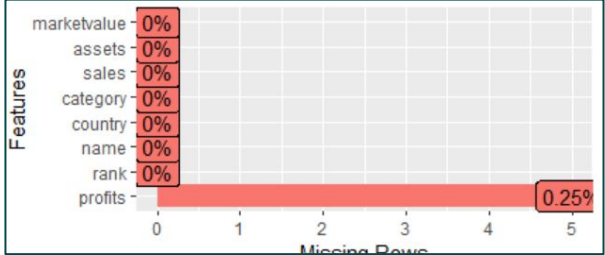
Menghitung Median dari Keuntungan Perusahaan di (1)US, (2)UK, France, dan Germany



Nomor 1.1

Menghitung Median dari Keuntungan Perusahaan di (1)US, (2)UK, France, dan Germany

1. Mengidentifikasi *Missing Value*

Syntax	Output
<p>Menghitung jumlah data yang kosong pada setiap kolom</p> <pre>colSums(is.na(forbes))</pre>	<pre>rank 0 name 0 country 0 category 0 sales 0 profits 5 assets 0 marketvalue 0</pre>
<p>Membuat grafik persentase <i>missing value</i> untuk menentukan proses perlakuan terhadap <i>missing value</i></p> <pre>plot_missing(forbes)</pre>	 <p>Missing value pada profits kurang dari 5% → melakukan omit</p>

Nomor 1.1

Menghitung Median dari Keuntungan Perusahaan di (1)US, (2)UK, France, dan Germany

2. Memproses *Missing Value*

Syntax	Output
<p>Membuat variabel baru yang sudah bersih dari data yang hilang</p> <pre>forbes2<-na.omit(forbes)</pre>	<p>Baris dengan data NA hilang</p>

Nomor 1.1

Menghitung Median dari Keuntungan Perusahaan di (1)US, (2)UK, France, dan Germany

3. Melakukan *filtering* data United States

Syntax	Output
<p>Membuat variabel baru yang berisi data perusahaan United States</p> <pre>forbes_us <- filter(forbes2, country == 'United States')</pre>	<p>Variabel baru: forbes_us</p>

4. Menghitung median data United States

Syntax	Output
<p>Membuat variabel baru yang perhitungan median dari kolom profits pada data forbes_us</p> <pre>median_us <- median(forbes_us\$profits)</pre>	<p>Nilai median</p> <pre>[1] 0.24</pre>

Nomor 1.1

Menghitung Median dari Keuntungan Perusahaan di (1)US, (2)UK, France, dan Germany

5. Melakukan *filtering* data UK, France, dan Germany

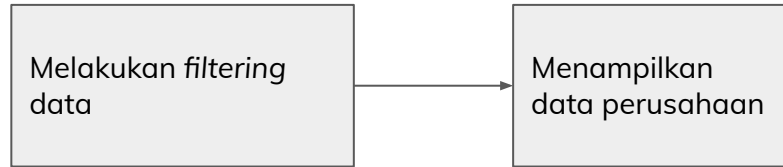
Syntax	Output
<p>Membuat variabel baru yang berisi data perusahaan UK, France, dan Germany</p> <pre>forbes_eu<- filter(forbes2, country == 'United Kingdom' country == 'France' country == 'Germany')</pre>	<p>Variabel baru: forbes_eu</p>

6. Menghitung median data United States

Syntax	Output
<p>Membuat variabel baru yang perhitungan median dari kolom profits pada data forbes_eu</p> <pre>median_eu <- median(forbes_eu\$profits)</pre>	<p>Nilai median</p> <pre>[1] 0.21</pre>

Nomor 1.2

Menentukan Perusahaan dari Jerman yang Memiliki Besar *Profit* Negatif



Nomor 1.2

Menentukan Perusahaan dari Jerman yang Memiliki Besar *Profit* Negatif

1. Melakukan *filtering* data

Syntax	Output
<p>Membuat variabel baru yang berisi data perusahaan dari Jerman yang nilai <i>profit</i>-nya negatif, atau di bawah 0</p> <pre>neg_profit <- filter(forbes2, country == "Germany" & profits < 0</pre>	<p>Variabel baru: neg_profit</p>

Nomor 1.2

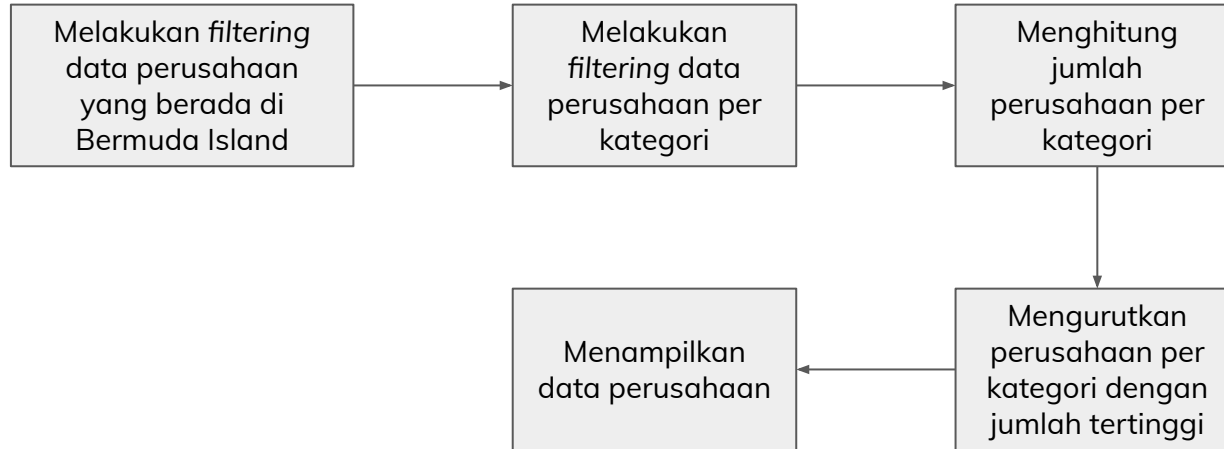
Menentukan Perusahaan dari Jerman yang Memiliki Besar *Profit* Negatif

2. Menampilkan data perusahaan

Syntax	Output
<p>Menampilkan nama perusahaan yang sesuai dengan kondisi</p> <pre>neg_profit\$name</pre>	<p>Terdapat 13 perusahaan di Jerman yang merugi sebagai berikut</p> <pre>"Allianz worldwide" "HVB-HypoVereinsbank" "BHW Holding" "mg technologies" "Mobilcom" "Deutsche Telekom" "Commerzbank" "Bankgesellschaft Berlin" "Nurnberger Beteiligungs" "E.ON" "Infineon Technologies" "W&W-Wustenrot" "SPAR Handels"</pre>

Nomor 1.3

Mencari kategori perusahaan yang paling banyak di Bermuda Island dan jumlahnya



Nomor 1.3

Mencari kategori perusahaan yang paling banyak di Bermuda Island dan jumlahnya

1. Melakukan *filtering* data

Syntax	Output
<p>Filter pertama dilakukan untuk perusahaan yang berada pada Bermuda Island</p> <pre>bermuda <- filter(forbes2, country == 'Bermuda')</pre>	Variabel baru: bermuda
<p>Filter didasarkan atas perusahaan per kategori yang berada di Bermuda Island, kemudian dihitung jumlahnya</p> <pre>bermuda2 <- bermuda %>% group_by(category) %>% summarise(count = n())</pre>	Variabel baru: bermuda2

Nomor 1.3

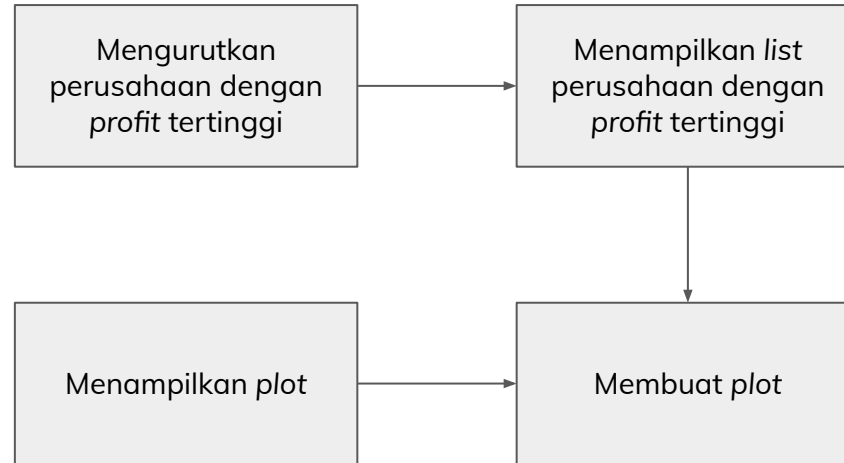
Mencari kategori perusahaan yang paling banyak di Bermuda Island dan jumlahnya

2. Menampilkan data bisnis yang paling banyak di Bermuda Island

Syntax	Output
<p>Membuat variabel baru yang merupakan kategori perusahaan di Bermuda Island yang diurutkan berdasarkan jumlah terbanyak</p> <pre>bermuda3 <- arrange (bermuda2, desc (count))</pre>	<p>Variabel baru: bermuda3</p>
<p>Menampilkan variabel baru</p> <pre>bermuda3</pre>	<p>Kategori perusahaan yang paling banyak di Bermuda Island adalah <i>insurance</i> dengan jumlah 10, disusul dengan Conglomerates dan Oil & gas operations dengan jumlah masing-masing adalah 2</p> <pre># A tibble: 9 x 2 category count <chr> <int> 1 Insurance 10 2 Conglomerates 2 3 Oil & gas operations 2 4 Banking 1 5 Capital goods 1 6 Food drink & tobacco 1 7 Food markets 1 8 Media 1 9 Software & services 1</pre>

Nomor 1.4

Membuat grafik untuk 20 perusahaan dengan *profit* tertinggi



Nomor 1.4

Membuat grafik untuk 20 perusahaan dengan *profit* tertinggi

1. Mengurutkan 20 bisnis dengan *profit* tertinggi

Syntax	Output																																																														
<p>Membuat variabel baru yang merupakan 20 bisnis dengan profit tertinggi diurutkan dari tinggi ke rendah</p> <pre>forbes20 <- forbes2 %>% arrange (desc(profits)) %>% slice_head(n=20)</pre>	<p>Variabel baru: forbes20</p>																																																														
<p>Menampilkan variabel baru dengan 6 profit tertinggi</p> <pre>head(forbes20)</pre>	<p>Tampilan list 6 bisnis dengan profit tertinggi. Didapatkan ExxonMobil sebagai perusahaan dengan profit tertinggi.</p> <table><tr><th>rank</th><th>name</th><th>country</th><th>category</th><th>sales</th><th>profits</th><th>assets</th><th>marketvalue</th></tr><tr><td>1</td><td>4</td><td>ExxonMobil</td><td>United States</td><td>Oil & gas operations</td><td>222.88</td><td>20.96</td><td>166.99</td><td>277.02</td></tr><tr><td>2</td><td>1</td><td>Citigroup</td><td>United States</td><td>Banking</td><td>94.71</td><td>17.85</td><td>1264.03</td><td>255.30</td></tr><tr><td>3</td><td>2</td><td>General Electric</td><td>United States</td><td>Conglomerates</td><td>134.19</td><td>15.59</td><td>626.93</td><td>328.54</td></tr><tr><td>4</td><td>6</td><td>Bank of America</td><td>United States</td><td>Banking</td><td>49.01</td><td>10.81</td><td>736.45</td><td>117.55</td></tr><tr><td>5</td><td>5</td><td>BP</td><td>United Kingdom</td><td>Oil & gas operations</td><td>232.57</td><td>10.27</td><td>177.57</td><td>173.54</td></tr><tr><td>6</td><td>20</td><td>Freddie Mac</td><td>United States</td><td>Diversified financials</td><td>46.26</td><td>10.09</td><td>752.25</td><td>44.25</td></tr></table>	rank	name	country	category	sales	profits	assets	marketvalue	1	4	ExxonMobil	United States	Oil & gas operations	222.88	20.96	166.99	277.02	2	1	Citigroup	United States	Banking	94.71	17.85	1264.03	255.30	3	2	General Electric	United States	Conglomerates	134.19	15.59	626.93	328.54	4	6	Bank of America	United States	Banking	49.01	10.81	736.45	117.55	5	5	BP	United Kingdom	Oil & gas operations	232.57	10.27	177.57	173.54	6	20	Freddie Mac	United States	Diversified financials	46.26	10.09	752.25	44.25
rank	name	country	category	sales	profits	assets	marketvalue																																																								
1	4	ExxonMobil	United States	Oil & gas operations	222.88	20.96	166.99	277.02																																																							
2	1	Citigroup	United States	Banking	94.71	17.85	1264.03	255.30																																																							
3	2	General Electric	United States	Conglomerates	134.19	15.59	626.93	328.54																																																							
4	6	Bank of America	United States	Banking	49.01	10.81	736.45	117.55																																																							
5	5	BP	United Kingdom	Oil & gas operations	232.57	10.27	177.57	173.54																																																							
6	20	Freddie Mac	United States	Diversified financials	46.26	10.09	752.25	44.25																																																							

Nomor 1.4

Membuat grafik untuk 20 perusahaan dengan *profit tertinggi*

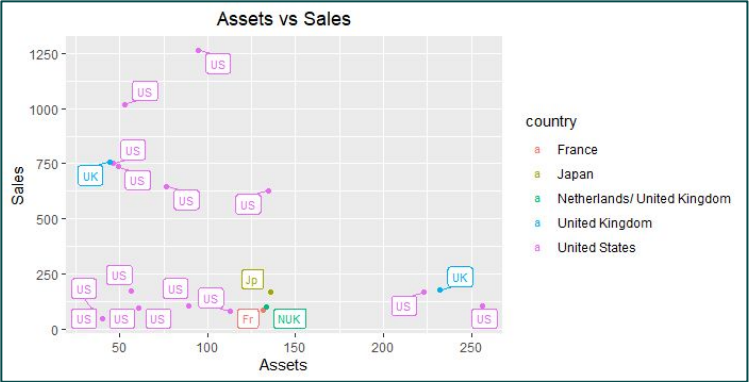
2. Membuat variabel untuk membuat grafik

Syntax	Output
<p>Membuat variabel baru yang merupakan scatter plot dari 20 perusahaan yang memiliki profit tertinggi dengan x adalah sales dan y adalah assets. Plot dirapikan serta diberi warna dan label per negara.</p> <pre>vis_forbes20 <- ggplot(data=forbes20, aes(x=sales, y=assets, color = country))+geom_jitter()+ geom_label_repel(aes(label=abbreviate(c ountry,2)), size=3)+ theme(plot.title = element_text(hjust = 0.5))</pre>	<p>Variabel baru: vis_forbes20</p>

Nomor 1.4

Membuat grafik untuk 20 perusahaan dengan *profit tertinggi*

3. Menampilkan visualisasi data ke dalam bentuk grafik

Syntax	Output
<p>Menampilkan plot dari variabel baru yang terbentuk</p> <pre>print(vis_forbes20+ labs(title = "Assets vs Sales", x = 'Assets', y = 'Sales'))</pre>	<p>Tampilan scatter plot 20 perusahaan dengan <i>profit tertinggi</i> yang nama dan warnanya dibedakan per negara.</p>  <p>country</p> <ul style="list-style-type: none">FranceJapanNetherlands/ United KingdomUnited KingdomUnited States

Nomor 1.5

Menentukan nilai rata-rata penjualan dan jumlah perusahaan yang memiliki keuntungan diatas 5 miliar US dollar pada setiap negara

1

Melakukan *grouping* data berdasarkan nama negara

Menghitung nilai rata-rata penjualan setiap negara

2

Melakukan *filtering* data

Memiliki keuntungan diatas 5 miliar?

Ya

Melakukan *grouping* data berdasarkan nama negara

Tidak

Hapus data

Hitung jumlah perusahaan untuk setiap negara

Nomor 1.5

Menentukan nilai rata-rata penjualan dan jumlah perusahaan yang memiliki keuntungan diatas 5 miliar US dollar pada setiap negara

1. Melakukan *grouping* data berdasarkan negara yang kemudian ditentukan nilai-rata-rata dari tiap negara tersebut

Syntax	Output
<p>Membuat variabel baru berisi negara dengan nilai rata-rata penjualannya</p> <pre>avg_sales ← forbes2%>% group_by(country)%>%summarise(mean(sales))</pre>	<p>Variabel baru: Total 61 negara dengan nilai rata-rata (mean) penjualan yang dimiliki</p> <pre>country `mean(sales)` <chr> <dbl> 1 Africa 6.82 2 Australia 5.24 3 Australia/ United kingdom 11.6 4 Austria 4.14 5 Bahamas 1.35 6 Belgium 10.1 7 Bermuda 6.84 8 Brazil 6.34 9 Canada 6.43 10 Cayman Islands 1.66 # ... with 51 more rows</pre>

Nomor 1.5

Menentukan nilai rata-rata penjualan dan jumlah perusahaan yang memiliki keuntungan diatas 5 miliar US dollar pada setiap negara

2. Melakukan *filtering* data sesuai keuntungan yang diminta, kemudian *grouping* data berdasarkan negara, dan terakhir dihitung berapa perusahaan yang masuk dalam kategori keuntungan tersebut untuk setiap negaranya

Syntax	Output																						
<p>Membuat variabel baru berisi hasil filter data negara yang memiliki keuntungan diatas 5 miliar. Kemudian dihitung jumlah perusahaan untuk masing-masing negara</p> <pre>Avglagi ← forbes2%>% filter(profits>5) %>%group_by(country) %>%summarise(count=n()) %>% arrange(desc(count))</pre>	<p>Variabel baru: Total 9 negara dengan jumlah perusahaan setiap negaranya yang memiliki keuntungan diatas 5 miliar</p> <table><thead><tr><th>country</th><th>count</th></tr><tr><th><chr></th><th><int></th></tr></thead><tbody><tr><td>1 United States</td><td>20</td></tr><tr><td>2 Switzerland</td><td>3</td></tr><tr><td>3 United Kingdom</td><td>3</td></tr><tr><td>4 China</td><td>1</td></tr><tr><td>5 France</td><td>1</td></tr><tr><td>6 Germany</td><td>1</td></tr><tr><td>7 Japan</td><td>1</td></tr><tr><td>8 Netherlands/ United Kingdom</td><td>1</td></tr><tr><td>9 South Korea</td><td>1</td></tr></tbody></table>	country	count	<chr>	<int>	1 United States	20	2 Switzerland	3	3 United Kingdom	3	4 China	1	5 France	1	6 Germany	1	7 Japan	1	8 Netherlands/ United Kingdom	1	9 South Korea	1
country	count																						
<chr>	<int>																						
1 United States	20																						
2 Switzerland	3																						
3 United Kingdom	3																						
4 China	1																						
5 France	1																						
6 Germany	1																						
7 Japan	1																						
8 Netherlands/ United Kingdom	1																						
9 South Korea	1																						