

T	P	R	S	D <sub>3</sub>
{	}	}	}	}

Garbage                      Col

- Extracting Information from our data.
- Visual representation.
- The link between dependent and independent variables.

- Just a Graphical representation of our data.
- **Cleaning, Wrangling**, and Exploring of our data.

Gold  
↓  
Skills → Gold product

1. Import Libraries.
2. Load Your Data.
3. Check your Dataset Columns
4. Statistically analyze your data.
5. Remove Null Rows.
6. Check Datatype (Change if necessary)
7. Rename Column (if necessary)
8. Change your data however you like
9. Correlation matrix
- ✓ 10. Graphical analysis: Data Visualization
- ✓ 11. Interpretation (Business Intelligence)

vulnerability of data

Day

H	T	P	R
1	2	3	4
5	6	7	8
9	10	11	12

Balance  $\geq 8$

50K  $\geq 11$

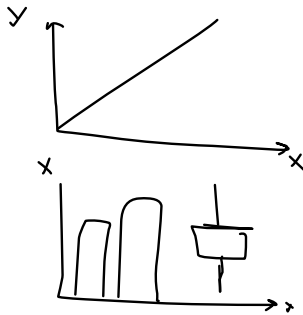
Utility  $\propto$  (only)

H T P R

8-10 20-30 9.01-9.05

2

1. Highly correlated  $> 0.5$
2. Medium Correlated  $> 0.3$  and  $< 0.5$
3. No correlation  $< 0.3$



How to decide which chart/graph to use?  
why?

$$\begin{matrix} M \\ 500 \end{matrix} \quad \begin{matrix} F \\ 300 \end{matrix} \Rightarrow 80\%$$

$$= \frac{500}{800} \times 100$$

male	2
not	3
male	4

male 9

$\sum_m |a|^2 = 50$   
 $2 \times \frac{50}{70} = 2 \times 0.71$

→ 500 -

\_\_\_\_\_

33333

\_\_\_\_\_

33333