

Data Visualization

Kathirmani Sukumar

Agenda

How to start visualizing any new data



What to look for in any type of chart



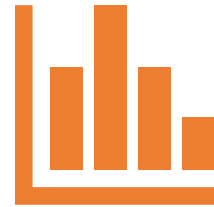
Doubts



Next steps



How do I ask
right
question?



Exploratory Data
Analysis



Domain Knowledge

Exploratory Data Analysis



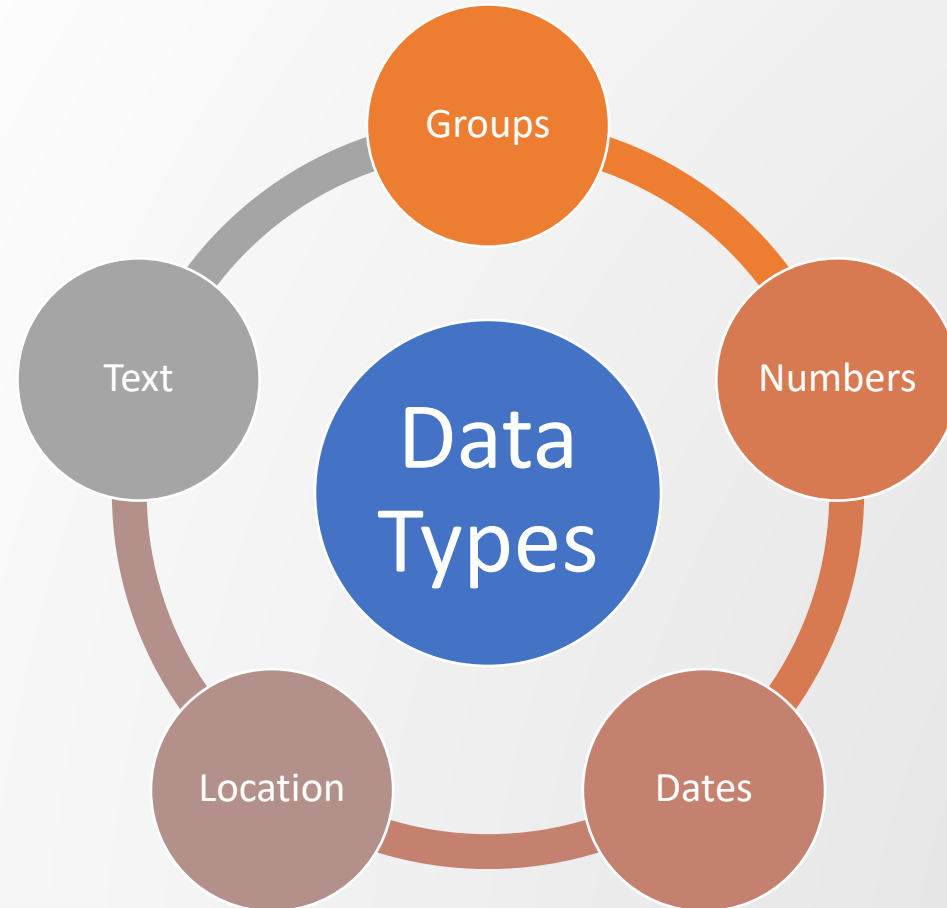
“

If you torture the data
long enough, it will
confess to anything.

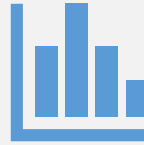
”

RONALD COASE

Data Types



Univariate



Groups

Frequency
analysis
(Pareto
principle)



Numbers

Distribution
(boxplot,
histogram)



Text

Frequency
analysis
(word cloud)

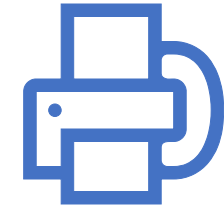
Bivariate Analysis

- Groups vs Numbers (Groupmeans analysis)
 - Bar charts, Box plots
- Numbers vs Numbers
 - Scatter Plots
 - Correlation analysis (Corrplot)
- Groups vs Groups
 - Crosstab analysis



Multi Variate Analysis

- 2 Groups, 1 Numbers
 - Treemap, Stacked Bars
- More than 2 Numbers
 - Scatter plot
- More than 2 Groups
 - Treemap, Sunburst



Trend Analysis

- Date and Number
 - Single/Multi Line chart
- Date, Number and Dimension
 - Stacked bar, Multiline chart
- Date and Text column
 - Multiple word clouds



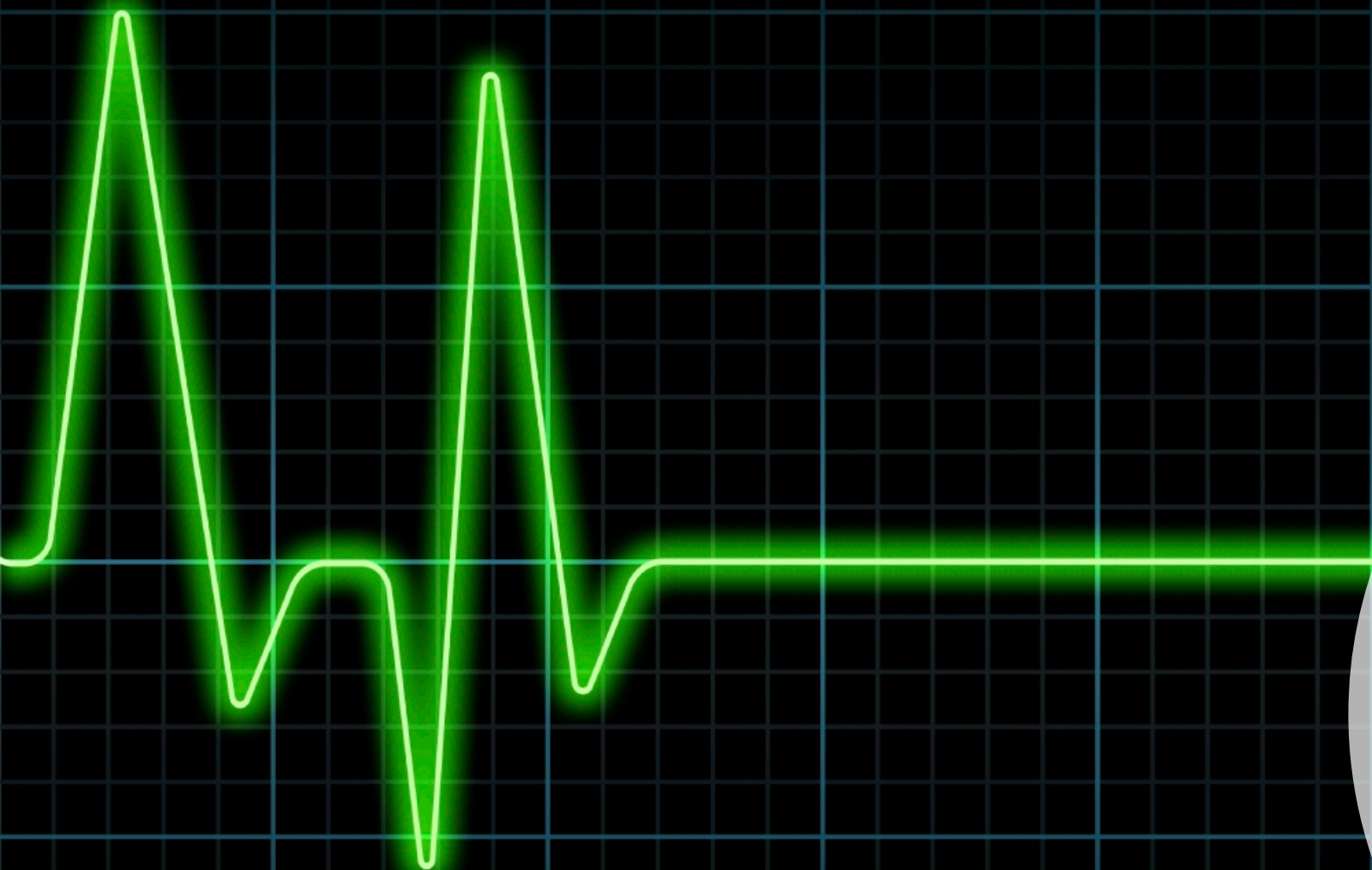


Hierarchical Charts

- Network graphs
- Sunburst
- Treemap
- Sankey
-

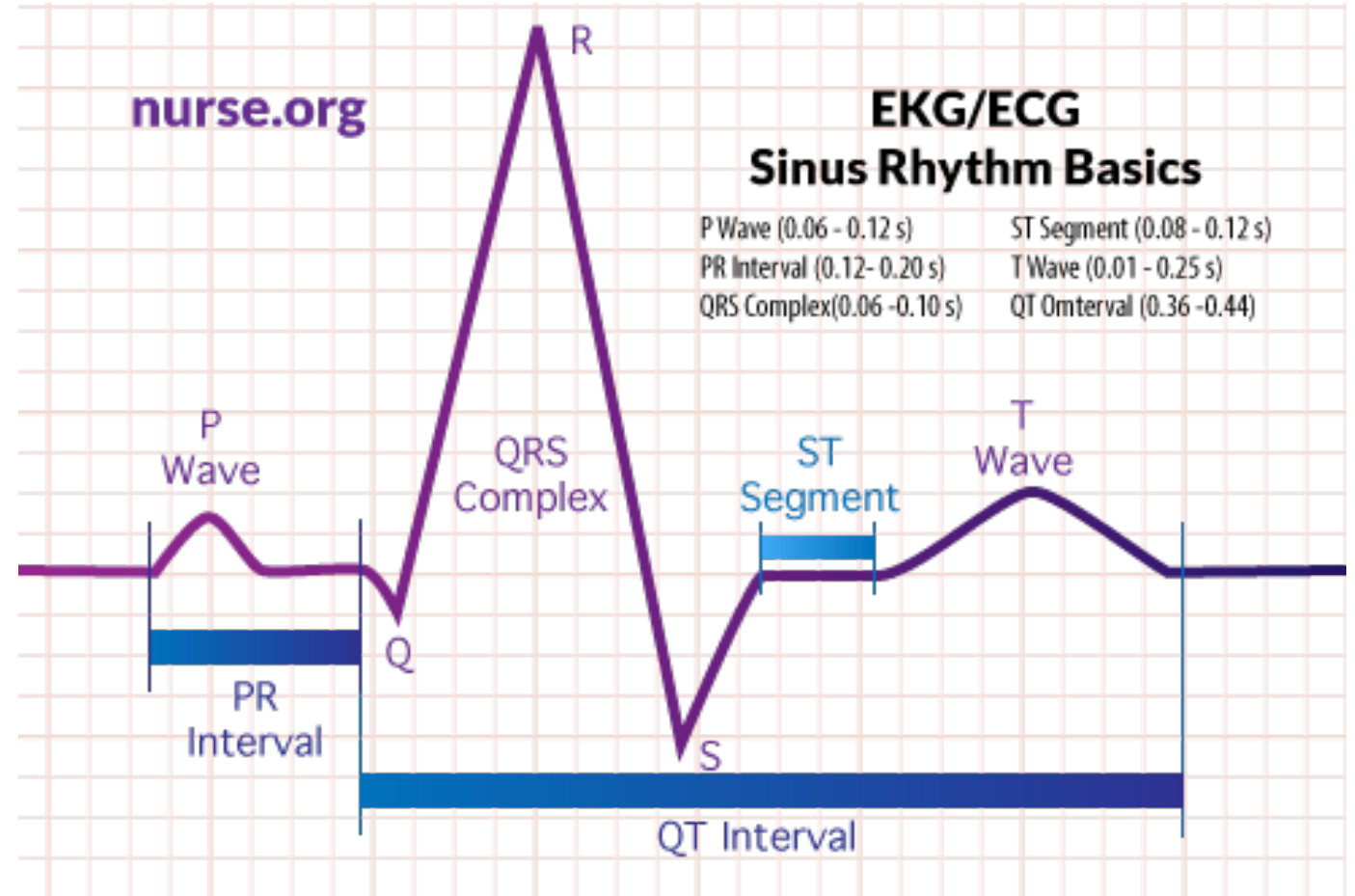


What do you look for
in any type of chart?

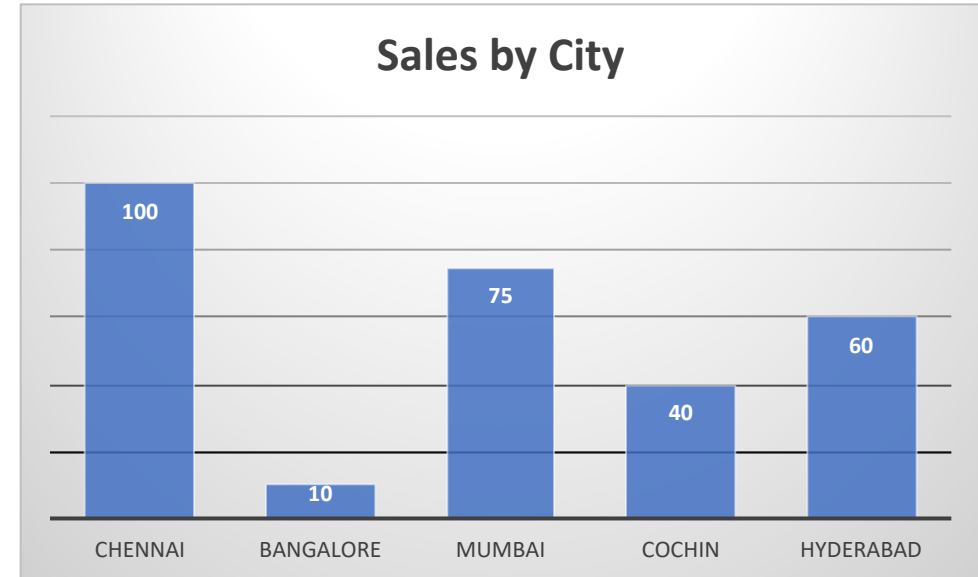
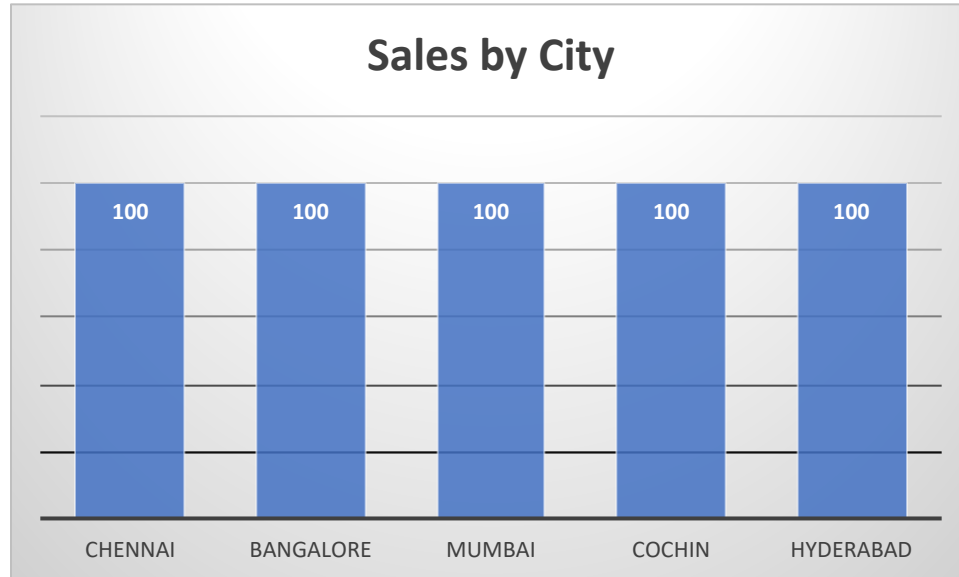


Is there anything
worth exploring
further???

Variance in data

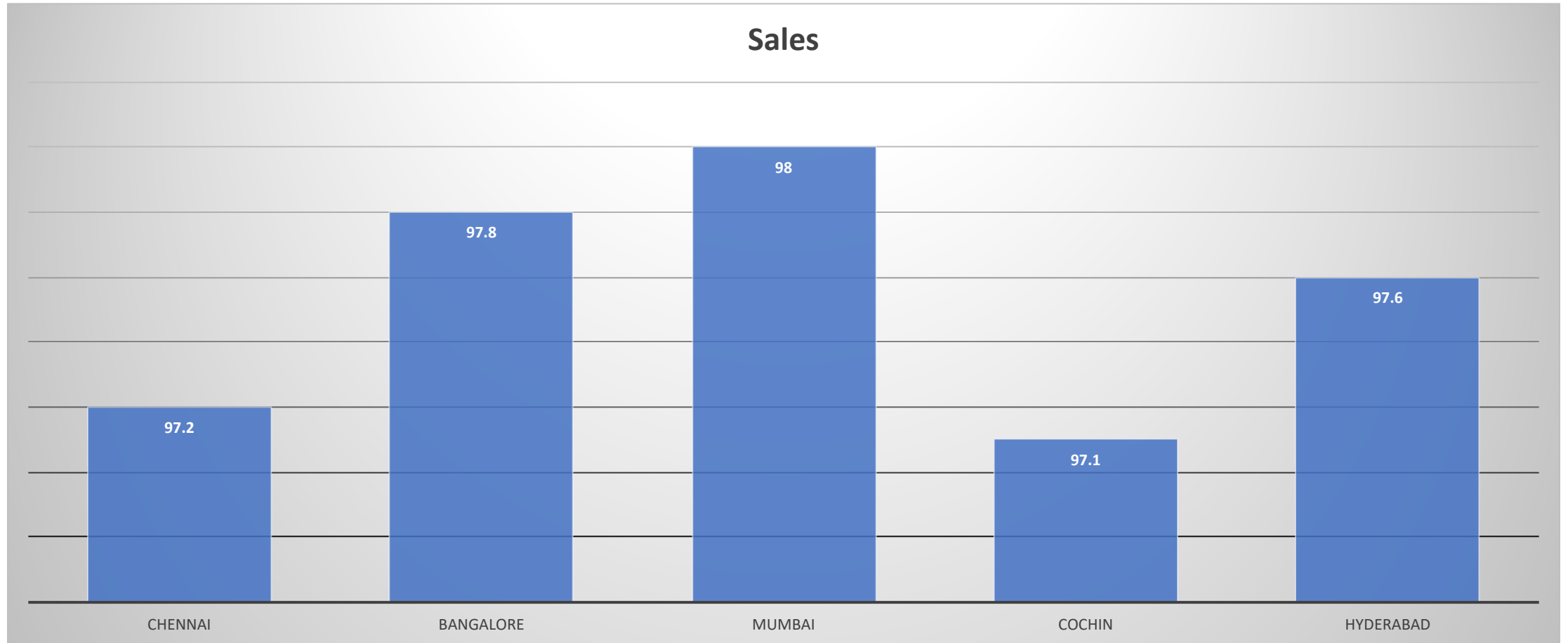


Bar chart



Which chart has some variance in it?

What about this?



Line chart



Seasonality



Anomalies



Trend

Scatter plot



Linearity



Groups



Outliers

Heatmap

- Single row/column darker/lighter
- In a row/column single cell alone darker/lighter
- Groups