BASIC GRAPH KNOWLEDGE

Bar charts
Used for comparing nominal entities.
Y-axis =quantity
X-axis= categorical or numerical
Vertical bars for less categories and horizontal (column chart) for more categories.

Column chart
Bar chart with 90 degrees rotated.
Used for nominal categories.

Histogram

Used to compare two metrics among several samples.
With probability distribution.
Difference btw hist and bar ?

Dual histogram
Used to compare males and females age on one histogram.

Stacked bar
Used to compare two metrics among several samples with actual numbers on Y-axis.Used to compare two metrics among several samples with percentages on Y-axis.
Here numbers is imp.
Eg: Travel rarely-(stayed /leaving),Travel frequently-(stayed /leaving).

100% stacked bar
Used to compare two metrics among several samples with percentages on Y-axis.
Here proportion is imp.
Eg: Travel rarely-(stayed (30%)/leaving70%),Travel frequently-(stayed(60%) /leaving(40%)).

Box plot
Used to show slices of a distribution.
Q1-25%,Q2- 50%(Median),Q3-75%.

Violin plot
Used to show several distribution completely all together.

Line chart
Used for temporal (time related)order.
Y-axis =quantity
X-axis= ordinal data or temporal.
See it as continues lines joining.

Pie chart
Used to represent proportions.
For categorical datatype.

Scatter plot
Used to compare two metrics among several samples.
Y-axis =numerical quantity
X-axis= numerical
See it as separate numbers or points.

Heat-map
Used to show a matrix of two features and its co-relation.

Area chart(like line chart)
Used to compare two histogram with area coloured.

Doughnut chart
Like a pie chart, but with hollow in the middle.Used for fancy visualisation.

Bubble chart
Scatter plot with a third variable for size .

Dendrograms
Used to show hierarchical clustering.
Eg: male,female with heights.

Radar or spider chart
Used to compare several metrics together instead of having many bars.
Eg: 5 students(aa, bb, cc, dd, ee) feedback on 4 features(clarity, energy, insight, assignments)

Waterfall chart
Used to show parts of a whole by adding or subtraction.
Eg:product and service revenue(420K& 210K).

Gantt chart
Used to show timelines.
X-axis= temporal
Y-axis= features

Pair-wise comparison chart
Used to two entities.

Survival curves (Kaplan-Meier)
Used to show failure occurs.
X-axis = temporal (time related)order.

CDF (Cumulative distribution function)
Used to compare two distributions.
Remember : Non-decreasing value and goes from 0 to 1 range.

Multiple graphs
Bar and line chart with dual Y-axis.
Helps to show two different quantities | units (profit(in %); revenue, net income(in $))on the same graph with temporal axis on X-axis.
Here scale both the Y-axises.

Mixed sub-chart
Pie chart and stacked chart
Show breakup of something using a pie and then take one component break it further using bar.
Eg: bananas(red, yellow),grapes(green, back),oranges (yellow, orange),apples(red, green, yellow)

ADDING INFO

Chart title - tells you about the chart.
Axis title - X-axis and Y-axis labels.
Axis units - for numerical inferences.
Grid lines - guiding lines ,should be light not too much fuzzy.If trend is only imp then no need to add it.
Legend - index column for chart.
Error bars - uncertainty in a column of grouped bar chart by (+,- std). Each bar saperately calculated.
Here it show that if data is less chance of error is also less vise versus.
Confidence intervals - fuzzy region of error for a feature(+,-).

BEAYTIFYING

Legibility of fonts - change on font size.
Consistency of fonts - Keep the font same of a chart.
Information to ink ratio - avoid more grid lines or grey background.
Use of coloured bars - use decent colours in chart.
Sorting - ordering of chart(alphabetically or numerical)

HIGHLIGHTING

Boxes - X-axis to be pointed at one category.
Call-outs and pointers - one point to be shown.
Specific colour for a bar.
Pie chart with a breakout.

WHAT TO USE

Dimension reduction tools for higher dimensional data
T-SNE

Data transformations
Log-scale
Used when rise of feature is exponential.Here order of the magnitude is imp.

Ratio of two variables
Used when variation within the variables is too large.

Polar form
Part of whole (pie, doughnut) or Radar chart or cyclical variables(hour, days, week, month, year)