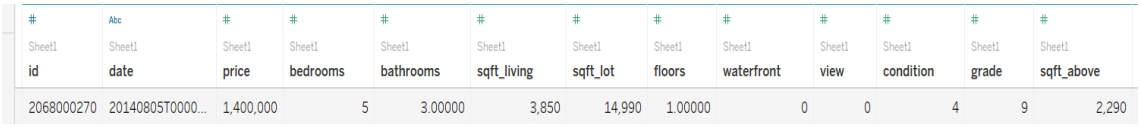
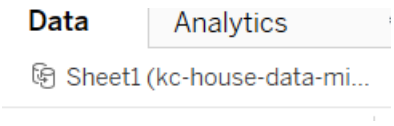
Tableau.

- Loading the file: Microsoft excel/text file
- Removing a file: right click dropdown and remove
- After loading the data just check with the data types of each variable. If any changes needed, update it.



Sheet1 id	Sheet1 date	Sheet1 price	Sheet1 bedrooms	Sheet1 bathrooms	Sheet1 sqft_living	Sheet1 sqft_lot	Sheet1 floors	Sheet1 waterfront	Sheet1 view	Sheet1 condition	Sheet1 grade	Sheet1 sqft_above
2068000270	20140805T0000...	1,400,000	5	3.00000	3,850	14,990	1.00000	0	0	4	9	2,290

- Change the connection to extract.



- Two cylinders represent extract connection.
- Sheet 1 and save
- Avoid loading two data sources simultaneously.
- Dimensions-categorical
- measures-numerical
- drag and drop pills that needs to compared.
- Row-y axis
- Column-x axis
- Add label, tooltip, sort data
- Standard to fit width to remove white space
- Show me icon helps to identify the other maps that could be build on the selected pills.
- Calculated fields:

  Analysis---create calculated fields--- name the column---perform calculations.

  Revenue=

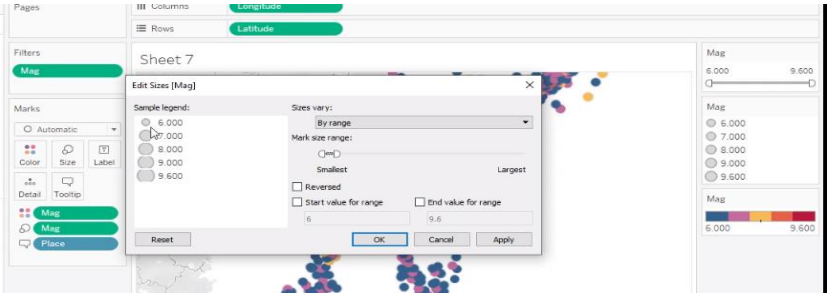
- Avoid making clutters instead make use of colors.
- Data table- double click on each pill.
- Date type variable- year and month. (line chart)
- Tree map-item, sales region, unit sold and revenue
- Dashboard- combine charts, hide title if needed.
- Filter each chart in the dashboard to get inference.
- Remove null:

  Analysis---create calculated fields--- name the column---perform calculations.

  If null ([column name],” others”)

- We can set three charts in one sheet with one column and three rows with sum, avg, medians.

- Analysis---create calculated fields--- name the column---perform calculations.
    - Worldwide gross/profitability ratio=budget
    - Worldwide gross-budget=net income
- Two measures-scatter plots. Remove aggregate measures.
- To get solid circles. Automatic---circles(beautification)
- Adding filters
    - Drag a pill to filters---pop up box---range---from to values will be there click ok
    - Then right click the pill in the filter---show filters---gilder will appear we can also set values. Set tool tip, size, color if needed.
- Export selected data from visualization
    - Select the required data point right click on any selected point and view data. Bottom---full data—then we export them.
- Scatter plots ---right click on chart trend lines---show trend lines
    - Edit trend lines-linear, log, exp
    - Desc trend lines
    - If scatter plots start above 6 or 5 axes, right click--- edit axis---change from automatic to fixed=6
- Latitude, longitude- geographical field
    - Double click on both---analysis---remove aggregate measures
    - Use search box to search for a country in the map
    - Pin icon to reset the zoomed map
    - Pan icon to pan through the map
    - Rectangular selection to select places
    - Add colors
    - Edit colors---Sunrise sunset---step colors---5---reverse— (choice based)
    - Add filters if needed
    - If we are applying filer and adding color on same dimension or measure, we should use center option,
    - Edit colors---advanced---center
- Edit sizes---change automatic to range and adjust the glider.



- Lasso selection in map- to select only the points needed
- Map---back ground maps---manage maps---map service pop up box---add---WMS server
  https://ows.terrestris.de/osm/service (free map service)
  - map—back ground map---Light (default map)
  - map layers---style---change mode, dark, satellite
  - we can enable cities in map layers
  - map layers---data layers at bottom--- US housing units, per capita income, population

- Background color---format---shading---worksheet---black
- Pages effect-time based data
  - Drop year to ---pages---add all numbers---we will get legend in right side---show history---select All---click play
- 49 unknows--- click---edit locations—country----from field---country---USA
- Pareto chart:
  - 80-20--- 80% of defects are contributed by 20% of items
  - 80% of money in India is handled by 20% of people
  - Column Sub-category, row- sum(sales)-bar chart, sum(sales)-line chart
  - sum(sales)-line chart- quick table calculation-running total- we need to change running total to percentage
  - second sum(sales)-edit table calculation-add secondary calculation—percentage of total
  - second sum(sales)-mark type-line chart
  - right click on second chart axis, dual axis—first sum(sales)—mark type—bar
  - enable tool tip if required
- Story:
  - requirements:
  - Slide 1 bar chart (basic chart)
  - Slide 2 bar chart (important comparative fields)
  - Dashboard 1-Line chart date + Slide 2 bar chart
  - Dashboard 2-Geographical wise + Slide 2 bar chart
  - Not mandatory to have dashboards in story charts can also be used.
  - Custom filters:
  - Add a pill to filters ---next---ok---show filters.
  - right click the pill---apply to worksheets---selected worksheets—choose the sheet you need to replicate the filter

- Story creation:
    - Story---new story
    - Select the field variable on which you want to build the story. Say for example, select table in dashboards as story captures that information
    - Size bottom -automatic
    - Add caption
    - Drag the image dashboard
    - Drag to add text
    - Next page—blank.
    - Drag all dashboards one by one.
- Inner join- drag the sheets. Data will repeat in tableau. When creating sets be cautious of what ID, you are using as it is a repeated feature
- Group data- select the bars of your choice and right click and third icon grouped members.
- Grouping cannot happen on date variables so convert in into numerical then do grouping
- Static set
    - 2015 data-double click ID,2015 Expenses. (Data table)
    - Least expensive company so sort in ascending order. We need top 10
    - Click First cell ---shift till 10^{th} cell--- right click—create set—name the set static 2015 expense--ok
    - Second set -Static 2015 profit (descending order)
    - Scatter plot od 2015 growth and 2015 revenue and add either static 2015 profit or static 2015 expense to filters
    - Based on employee size and industry try applying the sets to filters to make decisions.
- Dynamic set
    - ID---create---set---top/bottom---by field—10 choose field 2015 or 2014? name the set as dynamic set 2015 expense---ok
- Parameter control
    - To get 10,20, 30… glider type
    - Dimension---create parameter---(parameter expense) name it, data type, int based on the data type of the pill. Set it accordingly.
    - Allowable value- range min-10 max-50 step size=10
    - Go to dynamic 2015 expense set ---edit set—top----instead of 10 select the parameter created parameter expense.
    - Take the dynamic 2015 expense set to filters and click on the dynamic 2015 expense set under sets--- right click---show filters now the glider changes.
    - Condition parameter can also be set.
    - Make sure we use the same ID column while creating set.
    - Right click ID--- create set---condition--- by field--- 2015 expense---sum-- choose the symbol <= 2M
    - Right click ID--- create set---condition--- by field--- 2015 revenue---sum-- choose the symbol >= 10M
    - Right click on any one of the conditional dynamic set--- create combined set—best of both worlds---shared members in both sets---ok
    - Now drag the best of both worlds to filters

- Bullet charts target vs sales.
    - Swap reference line—right click –axis—swap reference line
    - To see the 60,70,80,90 target, right click—edit reference line—60,70—give values—change color—stoplight.
- Bubble chart.
- Box plot-aggregated box plot—use filter if you have more data i.e. more categories
- Wage diff=paid wage per year-prevailing wage per year
- Job title sub group vs wage diff—remove aggregate measure
- Now drag wage diff to filters and set the range to negative to find the under paid jobs.

R language:

- Rserve installation
- Rserve ()
- Help-settings and performance-manage external services connection—test connection—ok
- We can use R to perform calculations instead of calculated fields.
- Ex module 6.5 pdf

Data blending:

Sales happening daily basis but target fixed for monthly. So, we can join

Join should happen on unique columns only.

First two sheets are combined using ID column and click extract.

Load data icon---Microsoft excel-again load the sheet amazing market but now drag only the third data set. Now we get tow data sources

When we plot one pill from one source and other pill from other source, we get error. So we need to do data blending.

Data---edit relationships---custom—add

Order date--expand—select year from left side

Month order date—expand—select year from right side—ok

Again, click on add---expand order date—select month left side and month of order date—select month of order date from right side—click ok

Now when we click the sales source file, we can see a chain icon ensure that they are closed.

Now plot the pills from one source to another pill from another source.