

Tableau HW 2: Data Exploration and Chart Creation

Objectives of assignment: We will discuss how Tableau works with data and how it creates various charts in Part I and Part II. By the end of this assignment, you will understand

- Data uploading
- Data exploration: check the distribution of measures (outlier check)
- Chart creation

Tasks: Answer (6) multiple choice questions and upload a Tableau packaged workbook (.twbx) file that include (8) worksheets from Q6 -Q12

Deliverables: your answers to multiple choice questions in a Word doc and a .twbx file including (8) worksheets for Q6-Q12

Questions	Required Charts
Q6	A Box and Whisker plot
Q7	A bullet chart
Q8	A combo chart with bar and line marks
Q9	A dual axis bar chart
Q10	A map with viz in the tooltip (two worksheets needed - one map and one chart)
Q11	A dual axis map
Q12	Donut charts

HW data:

- Hospital revenue data.xls (provided on Canvas)
- World Indicators data (on Tableau)

Part I. Understanding Data with Tableau Desktop

- **HW data:** Hospital Revenue Data (Excel)
- **Deliverable:** A Word Document with your answers to Q1- Q6

Q1. First, go to **connect pane** and open an excel file named “revenue data.xlsx.” How many tables do you see on your left pane?

- 1) 3 sheets 2) 4 sheets 3) 5 sheets 4) 6 sheets

Next, let's connect all worksheets in the excel file. Make sure you click "data interpreter" before data join. Drag a fact sheet to **canvas and then**

- Merge with Fact table with fact.date key = dimDate.date key1 (***attention: you need to manually pick this PK-FK pair on Tableau**)
- Merge with dimEmployee with fact.emp key=dimEmployee.emp key
- Merge with dimFacility with fact.fac key= dimFacility.fac key
- Merge with dimPatient with fact.pat key= dimPatient.pat key
- Merge with dimProcedure with fact.proc key= dimProcedure.proc key.

This is the reason why a fact table has a composite key with pairs of PK-FK from different tables to link numeric variables to textual (descriptive) variable in dimensions tables.

Q2. To connect to multiple tables in a single data source at one time, what must be specified?

- 1) A blend
- 2) A calculation
- 3) A join
- 4) A hierarchy

Q3. (Circle one) "total charge", "procedure standard price", and "owed from patient" are dimensions (True or False)

Q4. When you connect to a data source, Tableau automatically separates date fields into hierarchies so you can easily break down the visualization. A hierarchy (aka tree structure) is a structure made up of two or more levels of related dimensions. For example, nation-state-county-city-zip code is an example of a hierarchy (reference: https://help.tableau.com/current/pro/desktop/en-us/qs_hierarchies.htm).

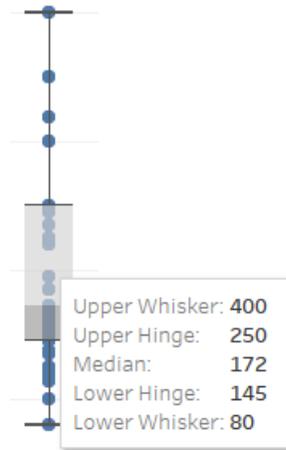
Location

- Country/Region
- State
- City
- Postal Code

Which of the following is NOT a hierarchy?

- 1) Time hierarchy: Year- quarter- month- week
- 2) Location hierarchy: USA-Florida-County-City
- 3) Sport gear: Gym – Kickboxing – Cycle- Tennis
- 4) Shopping: Clothing-Men's-Outer wear- Shirts

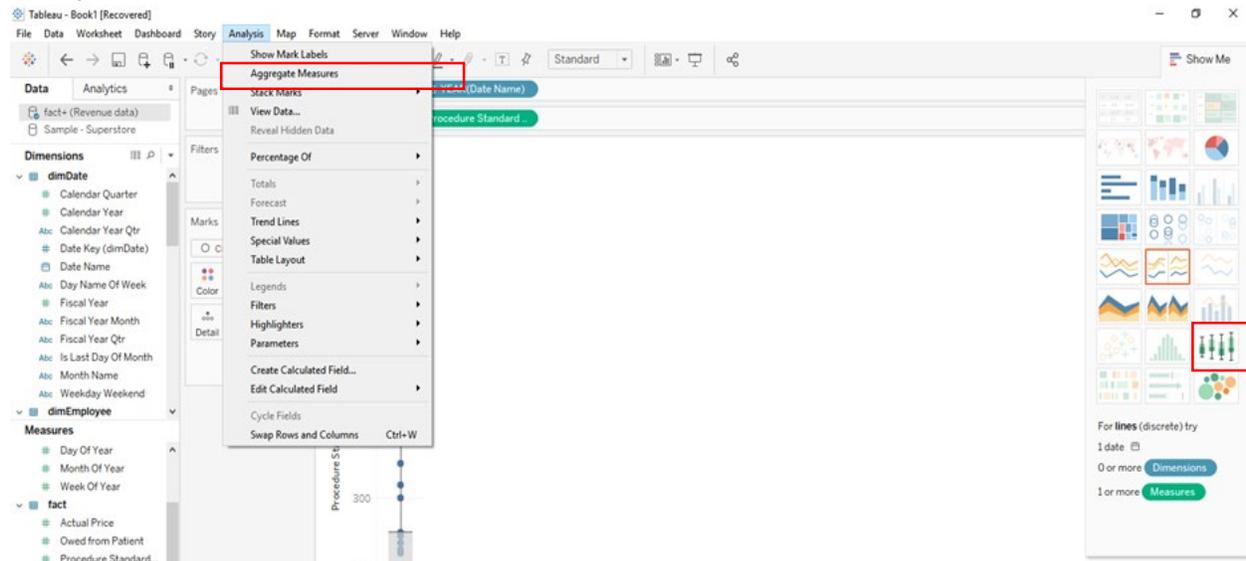
Q5. What does the box in a box plot above represent? Choose one (reference: https://help.tableau.com/current/pro/desktop/en-gb/buildexamples_boxplot.htm)



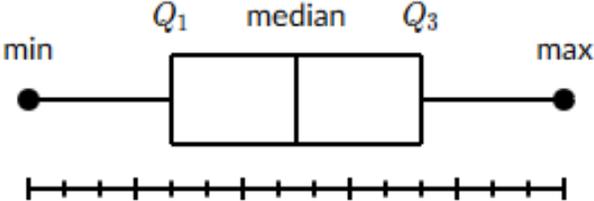
- 1) Maximum extent of the data
- 2) The range of the middle half of the data points
- 3) The median of the middle half of the data points
- 4) The outliers of the data

Q6. Draw a box and whisker plot using a dimension and a measure of your choice. To find the threshold for outliers, let's create a box-and-whisker plot on Tableau Desktop. Put **Procedure standard price** into Rows shelf, **Date Name** into Columns shelf, go to Analysis menu -> deselect "Aggregate Measures", then go to Show me panel and choose a "box and whisker plots." Name your worksheet as Q6.

Example



Typically, a box and whisker plot is used to check distribution of measures and detect outliers, using a five-number summary (Min-1st quartile (Q1)-2nd quartile (Q2, median) -3rd quartile (Q3)- Max). Outliers are less than $Q1 - 1.5IQR$ or greater than $Q3 + 1.5 IQR$, where Interquartile range (IQR) = $Q3 - Q1$ as shown below.



Then answer the following sub-questions:

Q6-1. First calculate IQR (Interquartile range)= $Q_3 - Q_1$.

Your answer:

Q6-2. Second, calculate minimum and maximum of thresholds.

Your answer:

$Q_1 - 1.5IQR =$
 $Q_3 + 1.5 IQR =$

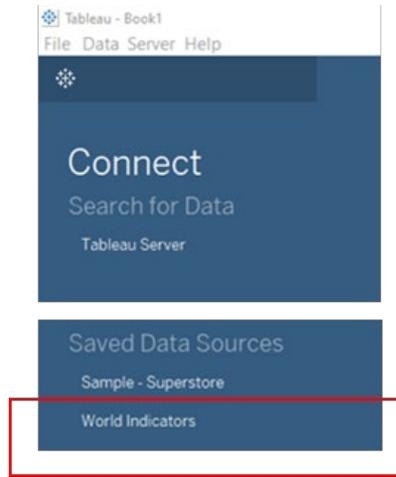
Q6-3. Make your conclusion as to whether there is any outlier in the data.

Your answer:

- Make sure you include the title of your box and whisker plot.
- In the Caption section, briefly describe your conclusion about outlier detection.

Part II. Chart Creation with Tableau Desktop

- **Homework data: World Indicators (under “Saved Data Sources” within Tableau Desktop)**



Using the dataset above, create the following charts using your own choice of dimension and measure from World Indicators data (for reference, you can both read the textbook chapter 3,6 and watch lecture recordings in Module 3 folder on Canvas).

- Make sure you include the title of your chart in each worksheet.
- In the Caption section, briefly describe your findings of your visualization in each worksheet.

Q7. A bullet chart

Q8. A combo chart with bar and line marks

Q9. A dual axis bar chart

Q10. A map with viz in the tooltip (two charts necessary, see textbook p.140-143)

Q11. A dual axis map

Q12. Donut charts