# Tableau Dashboard Guide

## **1. Analytics Team**

*- Team members:* Anh Quan Hua (Project Lead)

Nithini Bogahawattha

Xiaolu Li

Gouri Nandan Reddy Gangavaram

## **2. Tableau Dashboard Guide**

### A. Data Reception and Overview

Data Reception is the very first step for doing any data analysis is finding or collecting useful set of information/data and processing it before making any critical business decisions. The process of gathering, measuring, and evaluating various forms of information using a range of standardized, approved methodologies is also known as data collection. The basic goal of data collecting is to obtain trustworthy, information-rich data that can be used for analysis and important business decisions.

Data Overview or documentation is also a very important step which needs to be taken for having a proper understanding of the dataset including column descriptions, data types, and set of total number of records available and how many of them would be used for training, testing and for validation purposes.

After the Data Reception/Collection part, the next stage is the process of cleaning the data before doing any further processing of the data, because there might be several problems like challenges with poor data quality, and partial/null values which will affect the decision-making process and produce misleading outcomes which would affect decision making and damaging reputation.

Data cleaning tools: There are a lot of industry standard tools that are available for automating the cleaning process of different datasets and new data before further processing and evaluating it. Some of the tools are Tableau Prep, Dataloader.io, Datameer, Clear Analytics, etc. Python is a programming language which can be used coupled with a lot of already existing libraries from scikit learn, matplotlib, etc. to clean the dataset properly and transform and visualize the data further.

For data preparation, here are some of our recommendations:

The larger the dataset the better. More data can improve the accuracy of the analysis results.

The dataset needs to be tidy, and each feature value of every sample needs to be in a cell. If a cell contains multiple eigenvalues, the eigenvalues may not be separated during processing after importing into Tableau.

Null values may affect mathematical calculations of numerical variables, such as mean, median, etc. Nulls can also cause bugs in the early stages of processing the data. So, handle nulls as early as possible.

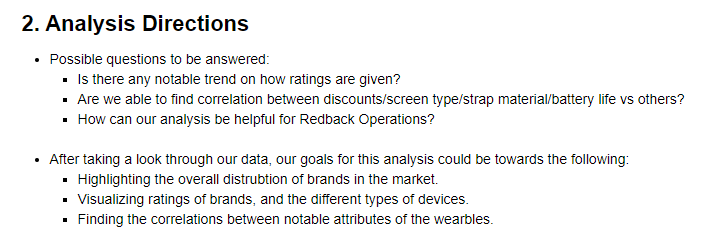
Outliers could be meaningful, so that we need to in-depth judgment. Some outliers are not caused by data collection errors and retaining outliers can make samples more abundant, if they are meaningful.

Simple relationship analysis could be performed before data is imported into Tableau. For example, use Python plots to check the relationship between variables. This helps us determine which variables to analyze. If there are no obvious relationships, then it may not be possible to analyze valid results in Tableau.

### B. Data analysis and visualization

In this section, we are going to look at how you should approach analyzing your datasets for the most optimal results with your Tableau Dashboard. The examples we use in the current part will be written in Python and utilize popular Data Science libraries to achieve our desired functionalities. Particularly, you can find the Jupyter Notebook for the examples in Redback’s Operation GitHub repository with the name “FINAL Fitness Tracker Sales Analysis.”

First, we want to remind ourselves of the goals and directions we want to take after our initial observations. Questions about the types of charts and graphs can be pulled from the analysis, the company’s purpose for the datasets and the possible trends that can be observed should be in your mind now. If you have not done that, now is a good time for you to list those out and write them in your notes or directly comment it in your coding file, see our example in Jupyter:

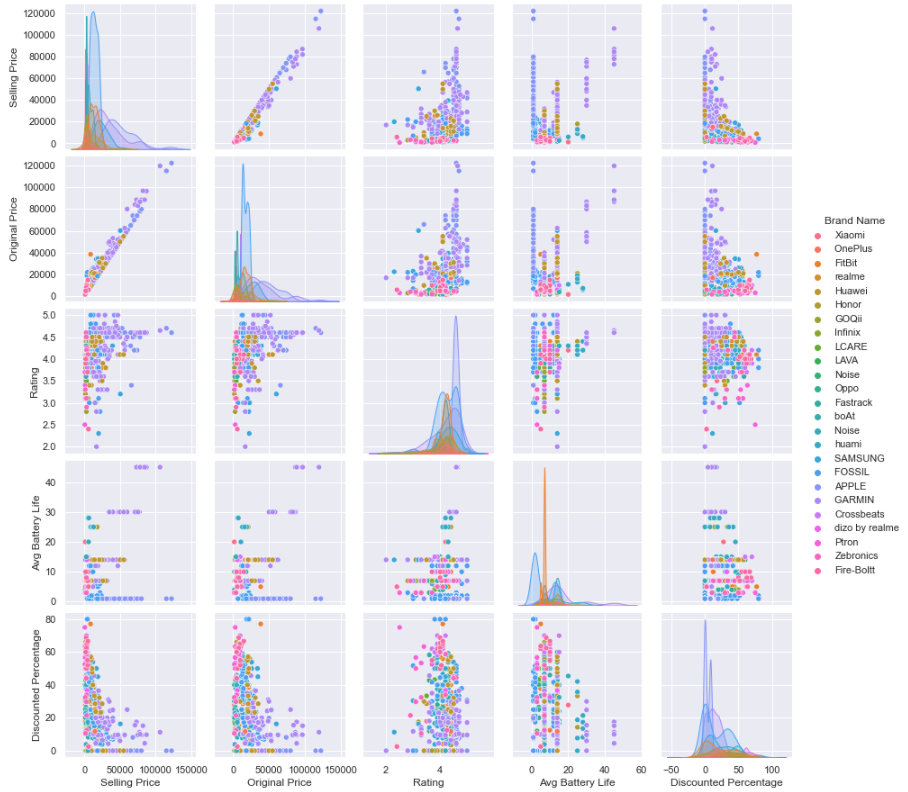


A good practice after cleaning and processing the data would be exporting your created “views” or newly cleaned dataset into a file. In Python, this can be done with:



Although Tableau has some powerful tools to process the data, this will help you to skip the redundant filtering process when creating visualizations later. This is ESPECIALLY important when you are transforming multiple tables and datasets in one to perform your analysis as Tableau does not support such intensive transformation of data.

Next, we want to start with the analysis. A helpful tip that works for us when starting this part is to begin with generalized visualizations that can provide some insights to your whole dataset. An example in Python would be a correlation heat map, or a set of pair plots for your whole data. The purpose of those is to help you immediately observe any direct correlation. Example here where we use the pairplot function to plot various graphs between pairs of attributes:



Finally, you would want to take note of the attributes used in your analysis’s visualizations. This means that once you are satisfied with a finding, a graph, or a trend after analyzing the datasets, it is good practice to know what data you are using – as this will help you fasten the process of producing the visualizations in Tableau later. On the ending note, visualizing as much as possible would be great too, since it helps you understand the rough form or how your graphs created in Tableau would be.

In summary, our guide for this part consists of the following:

- Be wary of your goals and purposes of the analysis – align your coding with them to help you focus on the right data.

- When creating complicated views or cleaning the datasets, remember to export them to a new file – this will be your direct data source used in Tableau, saving your time doing manual filtering and cleaning.

- Utilize generalized graphs that can work with almost any dataset to summarize key info.

- Take notes of the attributes you used for your desired graphs – this will help you choose the correct attributes and format them in Tableau (if needed).

- Do a lot of visualizing – this will serve as the basic sketches for your visualizations in Tableau.

### C. Tableau Dashboard Creation

For this section, to better understand the specifics when creating Tableau Dashboards – refer to Nithini Bogahawattha’s documentation for the Tableau Dashboard to better understand the context behind this.

The creation of the Tableau dashboard has got several steps included as follows.

* At first, I used the ‘Connect’ feature to a file ‘Excel’ in the left panel of Tableau software to open the ‘Fitness\_trackers\_Cleasned.xls.’
* Next to ensure to ‘Connect live’ I chose the ‘Live’ connection option and drag the ‘Data’ worksheet into the preview and then clicked on the ‘Sheet 1’ tab at the bottom of the screen.
* In ‘Sheet 1’ we will be able to identify our cleaned Categorical data as ‘Dimensions’ and Numerical data as ‘Measures’. This will make us easier to create the required Tableau workspace.
* While keeping a pause with the above three steps and before launching the dashboard I decided to define whom I will think will be the audience for the dashboard and what information they will be looking to explore and what insights they would like to gain through our live dashboard.

* I created a persona which would be the audience for the dashboard. Persona will include the following details of audience;

1. Who are they?
2. What is their skill level in using Data visualization?
3. What questions would they like to answer?
4. What action are we supporting?
5. What experience would we like to give them while using the dashboard?

* Moving on with my next steps I configured four workspaces. While dragging and dropping the ‘Dimensions’ and ‘Measures.’ These workspaces will be answering the following four questions for the audience;

1. Which display would be better when engaging in an intense workout?

       This graph has been created using categorical data ‘Brand Name’ and numerical data

       ‘Rating.’

1. What are the famous colours of the devices which are famous among the athletes?

To answer this question, I have used Categorical data ‘Brand Name’ and ‘Colour’ along with Numerical data ‘Ratings’ 

1. What is the most affordable Fitbit device according to Jeremy’s budget?

To create this graph, I have used Numerical data ‘Selling price’ and ‘Original price’ along with Numerical data ‘Brand name and ‘Model Name.’

1. What fitness tracker has the best average battery life as he is willing to wear the tracker throughout the day?

This graph was created using Categorical data ‘Brand Name’ and Numerical data ‘Number of reviews’ along with ‘Average life of a battery.’

* Moving on I configured the Tableau dashboard with ‘Add new Dashboard Icon’ at the bottom of the screen and set the Dashboard size to ‘Laptop’

* Next, drop and dragged the four worksheets on the Dashboard workspace in the place where it says, ‘Drop the sheets here.’
* Lastly, I published my Tableau dashboard to the public using the ‘Save to Tableau Public as’ function from the server menu to save the dashboard to the Tableau Public gallery. Along with this step, we will need to create a new profile to publish to the Tableau Public gallery.

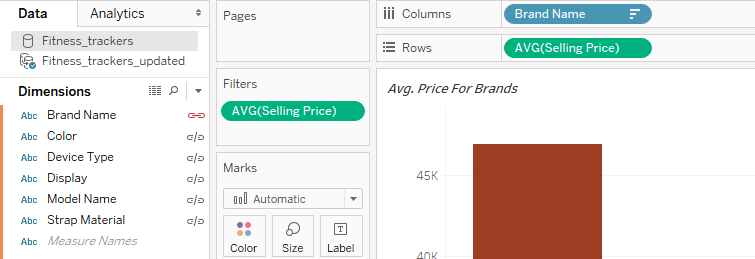
* During the whole process of creating my dashboard, I decided to document step by step. To identify and remember every step of mine which was used to create the live tableau dashboard. Documentation of the process has helped me in rechecking the process which I have followed.
* Also, to identify if I have missed any of the appropriate steps in creating the dashboard. This method could also help my other team members to identify if the dashboard is well established.
* As a conclusion, it could be said that the creation of the live dashboard will need much research, analysis, and upskilling activities. Which eventually led to a successful dashboard in the end.

### D. Using Existing Dashboards as Templates

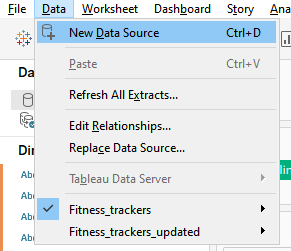
For future use cases, you might want to use some of the existing Tableau Dashboards as the base for your new Dash. If so, this is the part of our guide that you would like to take a look at.

Below are the steps to replacing the data source to utilize a Dashboard as a template. Examples are shown using Tableau 2019.2:

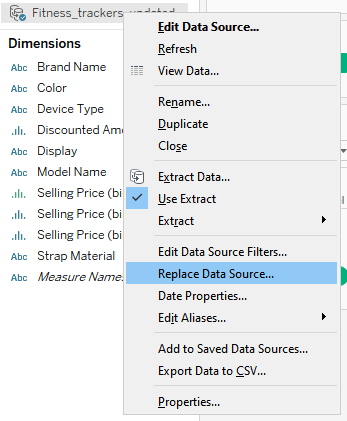
- First, in your desire worksheet, make sure that you have the new data source added under the “Data” Tab. As you can see, in this case we have “Fitness\_trackers” and “Fitness\_trackers\_updated”.



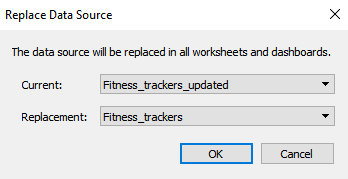
- In case you have yet to add the data source that you want to use instead of the old one, look for this button in the Tool Bar:  or go into this section above the navigation:



- Second, right-click on the current data source being connected and used:



- After clicking the “Replace Data Source” option, you should be greeted with the following windows where you can just choose the desired datasets:



## **3. Section Credits**

#### - A. Data Reception and Overview

*Xiaolu Li*

*Gouri Nandan Reddy Gangavaram*

#### - B. Data analysis and visualization

*Anh Quan Hua*

#### - C. Tableau Dashboard Creation

*Nithini Bogahawattha*

### - *D. Using Existing Dashboards as Templates*

*Anh Quan Hua*