Module 18 Challenge

Start Assignment

Due Thursday by 11:59pm **Points** 100 **Submitting** a text entry box or a website url

Background

Congratulations on your new job! As the new lead analyst for the New York Citi Bike ↗ program, you are now responsible for overseeing the largest bike-sharing program in the United States. In your new role, you will be expected to generate regular reports for city officials looking to publicize and improve the city program.

Since 2013, the Citi Bike program has implemented a robust infrastructure for collecting data on the program's utilization. Each month, bike data is collected, organized, and made public on the Citi Bike Data ↗ webpage.

However, while the data has been regularly updated, the team has yet to implement a dashboard or sophisticated reporting process. City officials have questions about the program, so your first task on the job is to build a set of data reports to provide the answers.

Before You Begin

- Save this assignment to your Tableau Public account rather than GitHub.
- If you haven't already, make sure to create a Tableau Public account ↗.
- The free tier of Tableau only lets you save to their public server. So, each time you save your file, it will be uploaded to your Tableau Public profile.
- You can load and continue working on the same workbook.
- When you finish your assignment, you will submit the URL to your Tableau Public workbook along with any additional files used in your analysis.

Files

You will be selecting files to download from the Citi Bike Data ↗ source, as described in the following instructions.

NOTE

Do not download every zip file available, as combining this much data will cause issues in Tableau. 1-3 zip files should be sufficient to meet the challenge requirements. Tableau restricts data file size to 1GB.

Note also that the data structure has changed over time, so it may be simpler for you to select data that uses the same columns across multiple files, preferably from within the past year.

Instructions

Your task in this assignment is to aggregate the data found in the Citi Bike Trip History Logs and find two unexpected phenomena.

1. Design 2–5 visualizations for each discovered phenomenon (4–10 total). You may work with a timespan of your choosing. Optionally, you can also merge multiple datasets from different periods.
   - The following are questions you may wish to answer. Do not limit yourself to these questions; they are suggestions for a starting point. Be creative!
     - How many trips have been recorded in total during the chosen period?
     - By what percentage has total ridership grown?
     - How have the proportions of short-term customers and annual subscribers changed?
     - What are the peak hours when bikes are used during the summer months?
     - What are the peak hours when bikes are used during the winter months?
     - Today, what are the top 10 stations in the city for starting a journey? Based on data, why do you hypothesize these are the top locations?
     - Today, what are the top 10 stations in the city for ending a journey? Based on data, why?
     - Today, what are the bottom 10 stations in the city for starting a journey? Based on data, why?
     - Today, what are the bottom 10 stations in the city for ending a journey? Based on data, why?
     - How does the average trip duration change by the type of user? (This may be under "User Type" or "member_casual" depending on the period the data is from).
     - What is the average distance in miles for a bike trip?
     - Which bikes (by ID) are most likely due for repair or inspection in the timespan?
     - How variable is the utilization by bike ID?
2. Use your visualizations (not necessarily all of them) to design a dashboard for each phenomenon. The dashboards should be accompanied by an analysis explaining why the phenomenon may be occurring.
3. Create one of the following visualizations for city officials:
   - **Basic:** A static map that plots all bike stations with a visual indication of the most popular locations to start and end a journey, with zip code data overlaid on top.
   - **Advanced:** A dynamic map that shows how each station's popularity changes over time (by month and year). Again, with zip code data overlaid on the map.
   - The map you choose should also be accompanied by a write-up describing any trends that were noticed during your analysis.
4. Create your final presentation:
   - Create a Tableau story that brings together the visualizations, requested maps, and dashboards.
   - Ensure your presentation is professional, logical, and visually appealing.

Considerations

Remember, the people reading your analysis will NOT be data analysts. Your audience will be city officials, public administrators, and heads of New York City municipal departments. Your data and analysis need to be presented in a way that is focused, concise, easy to understand, and visually compelling. Your visualizations should be colorful enough to be included in press releases, and your analysis should be thoughtful enough to inform programmatic changes.

Assessment

Your final product will be assessed on the following metrics:

- Analytic Rigor
- Readability
- Visual Appeal

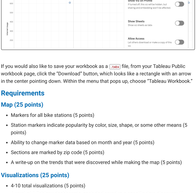
Hints

- You may need to get creative with how you combine each of the CSV files. Don't just assume Tableau is the right tool for the job. At this point, you have a wealth of technical skills and research abilities. Dig for an approach that works, and go with it.
- Don't assume that the CDV format hasn't changed since 2013. Subtle changes to the formats in any of your columns can interfere with your analysis. Ensure that your data is consistent and clean throughout your analysis. (Hint: Start and End Time change at some point in the history logs).
- Consider building your visualizations with small extracts of the data (like single files) before attempting to import the whole thing. What you will find is that importing all 20+ million records of data will create performance issues quickly. Welcome to "Big Data."
- While utilizing all of the data may seem like a nice power play, consider the time course in making your analysis. Is data from 2013 the most relevant for making bike replacement decisions today? Probably not. Don't let overwhelming data fool you. Ground your analysis in common sense.
- Remember, data alone doesn't answer anything. You will need to accompany your data visualizations with clear and directed answers and analysis.
- As is often the case, your clients are asking for a LOT of answers. Be considerate about their "need to know" and the importance of not "cramming in everything." Of course, answer each question, but do so in a way that is organized and presentable.
- Since this is a project for the city, spend the appropriate time thinking through decisions on color schemes, fonts, and visual storytelling. The Citi Bike program has a clear visual style. As a suggestion, look for ways to have your data visualizations match their aesthetic.
- Pay attention to labels. What exactly is "time duration?" What's the value of "age of birth?" You will almost certainly need calculated fields to get what you need.
- Look for obvious outliers or false data. Not everyone who signs up for the program is answering honestly.
- In answering the question of "why" a phenomenon is occurring, consider adding other pieces of information, like socioeconomic or geographic data. Tableau has a map "layer" feature that you may find helpful.
- Don't be afraid to manipulate your data and play with settings in Tableau. Tableau is meant to be explored. We haven't covered everything that you'll need—so keep an eye out for new tricks!
- Treat this as a serious endeavor! This is an opportunity to show future employers that you have what it takes to be a top-notch analyst.

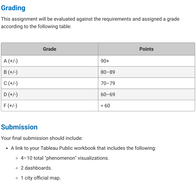
Sharing Your Work

To share your work, save your workbook to Tableau Public and submit the URL so that your TAs can grade it.

Save by using shortcuts (CTRL + S or CMD + S), or by clicking "File" then "Save to Tableau Public."



After saving, the visualization will be uploaded to your profile, which will automatically open in a new browser window. Adjust the settings of your new visualization by clicking the gear on the toolbar and toggling where to show the visualization and to allow access.



If you would also like to save your workbook as a .twbx file, from your Tableau Public workbook page, click the "Download" button, which looks like a rectangle with an arrow in the center pointing down. Within the menu that pops up, choose "Tableau Workbook."

Requirements

Map (25 points)

- Markers for all bike stations (5 points)
- Station markers indicate popularity by color, size, shape, or some other means (5 points)
- Ability to change marker data based on month and year (5 points)
- Sections are marked by zip code (5 points)
- A write-up on the trends that were discovered while making the map (5 points)

Visualizations (25 points)

- 4-10 total visualizations (5 points)
- A total of 2 Tableau dashboards, each dedicated to a specific data discovery (5 points)
- Dashboards are named appropriately (5 points)
- Data is cleaned such that data entry errors are removed and columns are correctly typed (5 points)
- Visualizations can logically be used to explore the data (5 points)

Tableau Story (25 points)

- Individual visualizations are used (3 points)
- Dashboards are used (5 points)
- A map is used (5 points)
- Visualizations on the same page are clearly related to one another (5 points)
- The story is informative and easy to navigate (5 points)

Analysis (25 points)

- Analysis is written in a markdown file or included in the Tableau Public workbook (5 points)
- Analysis describes the dashboards and any interesting data discoveries contained within them (5 points)
- Analysis on the chosen city official requested map detailing any noticeable trends (5 points)
- The written analysis references specific visualizations and interactive features (5 points)
- The document is written in a manner that a non-technical reader could understand (5 points).

Grading

This assignment will be evaluated against the requirements and assigned a grade according to the following table:

Grade	Points
A (+/-)	90+
B (+/-)	80–89
C (+/-)	70–79
D (+/-)	60–69
F (+/-)	< 60

Submission

Your final submission should include:

- A link to your Tableau Public workbook that includes the following:
  - 4–10 total "phenomenon" visualizations.
  - 2 dashboards.
  - 1 city official map.
  - 1 story.
  - A text or markdown file with your analysis of the phenomena you uncovered in the data.

To submit your Challenge assignment, click Submit, and then provide the URL to your Tableau Public workbook along with any additional files used in your analysis.

NOTE

You are allowed to miss up to two Challenge assignments and still earn your certificate. If you complete all Challenge assignments, your lowest two grades will be dropped. If you wish to skip this assignment, click Next, and move on to the next Module.

Comments are disabled for graded submissions in BootCamp Spot. If you have questions about your feedback, please notify your instructional staff or your Student Success Manager. If you would like to resubmit your work for an additional review, you can use the Resubmit Assignment button to upload new links. You may resubmit up to three times for a total of four submissions.

IMPORTANT

It is your responsibility to include a note in the README section of your repo specifying code source and its location within your repo. This applies if you have worked with a peer on an assignment, used code in which you did not author or create sourced from a forum such as Stack Overflow, or you received code outside curriculum content from support staff such as an Instructor, TA, Tutor, or Learning Assistant. This will provide visibility to grading staff of your circumstance in order to avoid flagging your work as plagiarized.

If you are struggling with a Challenge or any aspect of the curriculum, please remember that there are student support services available for you:

1. Office hours facilitated by your TA(s)
2. Tutor sessions (sign up ↗)
3. Ask the class Slack channel/get peer support
4. AskBCS Learning Assistants

© 2023 edX Boot Camp LLC