



## Unit 2 Project: Mozilla Firefox

---

### PROJECT OVERVIEW

It's 2010, and you are a junior product manager working on Mozilla Firefox.

You are chatting with an opinionated colleague of yours, and you ask her thoughts about your idea to revamp the bookmarks experience. She says that bookmarks aren't interesting and you should help her in her quest to improve support for having a large number of tabs. Your boss has asked you to put together a presentation recommending an area of focus for the next version of Firefox.

After considering the Data Analytics Workflow, you decide to look at a [recent test flight](#) with customer surveys and usage analytics to better inform your thinking. After that, you will decide which of the two areas to recommend.

*\*Note: The data required for this project can be queried from the AWS server you have been using in class, but please read through the “recent test flight” link above for the official documentation.*

---

### DELIVERABLES AND TIMELINE

#### 1. Initial Draft of Findings

Due: beginning of Lesson 12

Format: a four- to five-paragraph e-mail to your boss

- Describe the users represented in the data. Do you think they're representative of the population of Firefox users?
- Describe any initial findings related to your research questions.
- Describe what you still need to do for your presentation in order to make a sound recommendation.

#### 2. Five- to 10-minute Presentation:

Due: presented during Lesson 14

Format: Google Slides or PDF (Keynote/PPT need to be exported)

- This will be a final presentation to your boss and team at work. It will build on your earlier e-mail and summarize your overall findings and recommendation.

- Summarize the data in your presentation:
  - Present your findings, results, and recommendations to the class.
  - Review the Users table. Summarize the users represented in the survey.
  - Review the Survey table. How many of the total users completed the survey?
  - Of users who completed the survey, identify the number of users who are new to Firefox, as well as those who are longtime users.
  - Include other information that is important to your initial findings, and which may have affected your recommendation and final presentation.
- Summarize your exploration of the bookmark-usage feature in your presentation. Some sample questions you could answer include:
  - What's the median number of bookmarks? What's the average number?
  - What fraction of users launched at least one bookmark during the sample week?
  - What fraction of users created new bookmarks?
  - What's the distribution of how often bookmarks are used?
  - How does number of bookmarks correlate with how long the user has been using Firefox?
- Summarize your exploration of usage of browser tabs in the presentation. Some sample questions you could answer include:
  - What's the distribution of the maximum number of tabs?
  - Are there users who regularly have more than 10 tabs open?
  - What fraction of user have ever had more than five tabs open? What fraction of users have ever had more than 10 tabs open? What fraction of users have had more than 15 tabs open?
- Compare the two features and create a coherent recommendation for your boss.
  - Do you agree or disagree with your opinionated colleague?
  - Do you think the team should go in a different direction?
- Include analysis of data with statistics and outliers.
- Include recommendation with predicted outcomes and identified action items.
- Present any limitations and assumptions.
- Identify follow-up problems and questions for future analysis.

### 3. SQL Queries + Questions:

Due: beginning of Lesson 14

Format: .sql file

- List the appropriate business questions you were trying to answer.
- Use correct, efficient syntax for the functions and operators needed.
- Include corresponding output listed in the .sql file.

## GETTING STARTED

Consider the Data Analytics Workflow to guide you through each step:

- **Identify and Understand the Problem:** After considering the prompt, think generally about the problem and how you might want to approach it. Ideally, what type of data would you have to make a decision and produce a recommendation?
- **Access and Obtain the Data:** Your company, Mozilla, is running its own PostgreSQL server, and the engineers who designed the survey have added that data to three tables: events, survey, and users. Connect to the appropriate server using the software of your choice (pgAdmin III or DBVisualizer, for example).
  - To connect to the server, please use the following login credential with the `firefox_database`:
    - **url/host:**
      - West-coast (includes Austin and Chicago) - [analyticsga-cuwj8wuu6wbh.us-west-2.rds.amazonaws.com](https://analyticsga-cuwj8wuu6wbh.us-west-2.rds.amazonaws.com)
      - East-coast - [analyticsga-east2.c2ogkj5cvu3l.us-east-1.rds.amazonaws.com](https://analyticsga-east2.c2ogkj5cvu3l.us-east-1.rds.amazonaws.com)
      - APAC - [analyticsga-apac2.csuojbfcexhv.ap-southeast-2.rds.amazonaws.com](https://analyticsga-apac2.csuojbfcexhv.ap-southeast-2.rds.amazonaws.com)
      - Europe - [analyticsga-west1.c8yztvlaoku1.eu-west-1.rds.amazonaws.com](https://analyticsga-west1.c8yztvlaoku1.eu-west-1.rds.amazonaws.com)
    - **port:** 5432
    - **username:** analytics\_student
    - **password:** analyticsga
- **Understand the Data:** Your colleagues in the engineering department are very organized and have written documentation for the product managers who need to use their databases. Make sure you read it! [Link to documentation](#)—*note that the live site has been taken down but the web archive works great!*
  - **HINT:** Make sure to identify which user actions will be of most use to your research questions. Browse the tables and ensure you know how the data is organized and how the tables relate to each other. Answer basic questions about the data. When presenting the results of survey data, it's often good to inform your audience about the characteristics of the surveyed users. Do you think these users are representative of all Firefox users? Why or why not? What other limitations might exist with this data?
- **Prepare, Structure, and Clean the Data:** In order to be able to settle on your recommendation and create your presentation, you will need to write a number of queries to prepare, structure, filter, clean, and aggregate your data.

- **Analyze and Visualize the Data:** In order to determine findings and create a recommendation, you will want to use statistical (mean, median, mode, min, max, etc.) functions as well as other aggregate functions (count) to summarize your data. Export your data to Excel to create any visualizations that will help you either explore your data or provide evidence for your recommendation.
- **Present the Results of Analysis:** Based on your findings, create a coherent argument about which project is more worthwhile: tabs or bookmarks.
  - Create a PowerPoint presentation that summarizes your findings and makes a recommendation.
  - Make sure to include the results of your analysis and any visualizations that help boost your argument.

---

## RESOURCES

- Review specific class decks on how to use SQL and access the database.
- View the dataset [documentation](#).
- View the recent [Test Flight](#) data.
- Reference the Data Mining [Framework](#).
- Reference General Assembly's Data Analytics Workflow.

---

## EVALUATION

- Your presentation will be graded using the requirements section above as a rubric.
- The Excel worksheet containing the cleaned data will be evaluated using the requirements section above as a rubric.