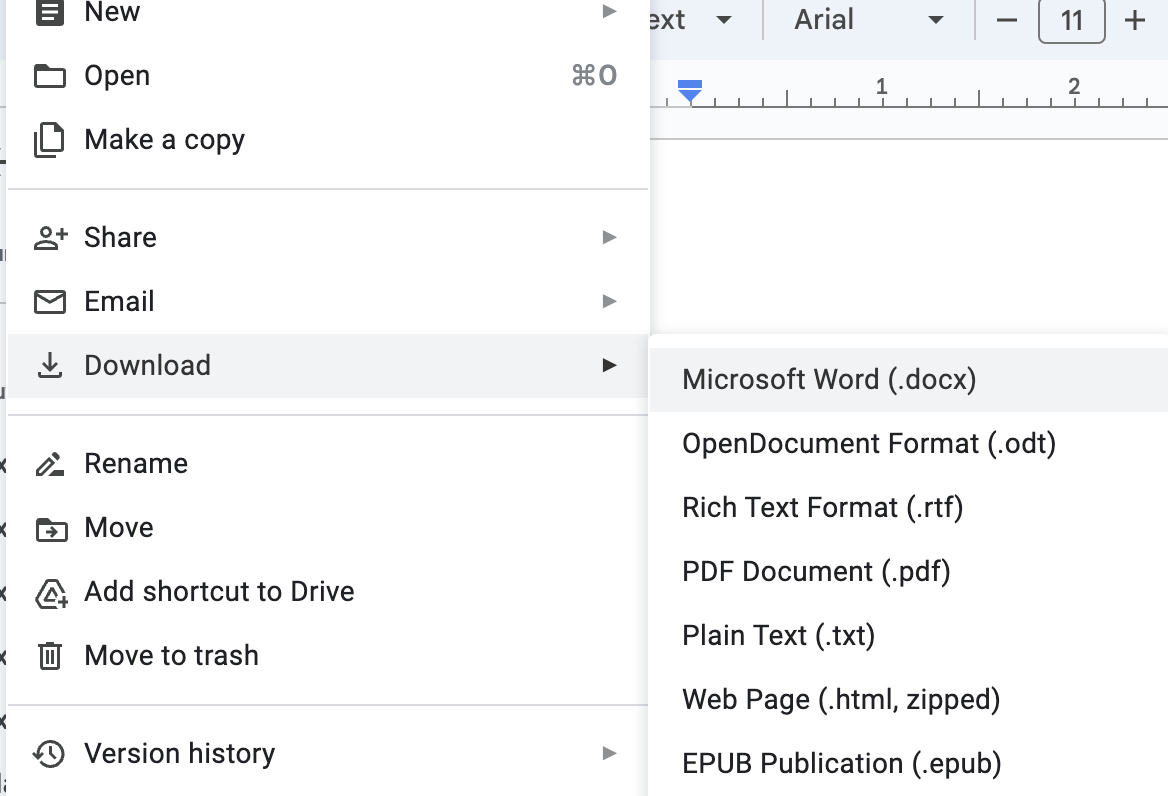
For this in-class activity, we will be working in groups of two or three.

Download a copy of this document by clicking on “File” → “Download” → “Microsoft Word (.docx)” in the menu above. You will use this doc to keep track of your work. Feel free to share the doc with teammates using Microsoft 365. You can also share with your team by making a copy of this file through google docs.



**Note**: text in italics should be edited by your group!

**Note**: Make sure exactly one team member submits your in-class activity document [here](https://forms.gle/TdJ3m38QVYSkXWpX8) *before* you leave lecture today!

**Exercise 1: Review the Dataset and Prototype Some Earnest Visualizations (20 mins)**

In this exercise, we will explore the provided dataset by creating some *earnest* visualizations. We’re not trying to fool anyone yet! We’re just trying to understand the data. The dataset we are using is from a visualization created by the Women’s Bureau, US Department of Labor. A google sheet of the data is available [here](https://docs.google.com/spreadsheets/d/1DBx70W6oKvpSXMUS0btyNeJ_PHiv7X0YLyCLUpvkNLo/view). A CSV file of the data is available [here](https://drive.google.com/file/d/1qQOUYkNgSvVpuqdmelFpFK5a_9-8B543/view). The original visualization is [here](https://www.dol.gov/agencies/wb/data/earnings/wage-gap-race-occupation).

Here are the fields of the sample dataset and some sample rows. Think about the data types (N, O, Q) and whether you might use them as dimensions or measures in your visualizations.

* **Sex**
* **Race and ethnicity:** Reported Race according to US Census data (please see the notes from [the original visualization](https://www.dol.gov/agencies/wb/data/earnings/wage-gap-race-occupation) for how race was recorded).
* **Median Earnings (USD):** Median annual earnings for full-time, year-round workers in the US.
* **Occupation Group:** The types of jobs measured

|  |  |  |  |
| --- | --- | --- | --- |
| **Sex** | **Race and ethnicity** | **Median Earnings (USD)** | **Occupation Group** |
| Men | All | 61,199 | All Groups |
| Women | All | 51,216 | All Groups |
| Men | Black | 70,968 | Management, business, and financial |
| Women | Black | 65,553 | Management, business, and financial |
| Men | Asian | 50,891 | Nat resources and construction |
| Women | Asian | N/A | Nat resources and construction |

Is there anything confusing about this dataset? Anything that doesn’t make sense? Please feel free to ask questions and get help during the exercise!

Now, the group should use whatever tools they want to create at least two earnest visualizations. Paste images of the different visualizations you come up with below.

*A graph with numbers and text

Description automatically generated*

What is the highest median earnings among different occupation groups?

*A graph of blue and orange bars

Description automatically generated*

We will do a share out towards the end of this exercise, to see what visualizations people came up with!

**Note**: Make sure exactly one team member submits your in-class activity document [here](https://forms.gle/TdJ3m38QVYSkXWpX8) *before* you leave lecture today!

**Exercise 2: Create Deceptive Visualizations (25 mins)**

Now, the group should create at least one *deceptive* visualization. The Deceptive Visualization lecture listed four types of deceptive visualizations. Consider these visualization types as you design your deceptive visualizations.

* Incorrect Visualizations
* Illegible Visualizations
* Bullshit Visualizations
* Unconventional Visualizations

Paste images of the different visualizations you come up with below.

A pie chart with text

Description automatically generatedA pie chart with text

Description automatically generated

Why do you think this visualization is deceptive (if you created more than one, state which visualization you picked)? Please explain your reasoning below. For example, what techniques did you use to make it deceptive?

*We plot the median earnings by using data in the category “Office and admin support” and see the median earnings between men and women among different race and ethnicity. The bar chart shows men earn more than women among different race and ethnicity, but we cannot tell this information from the pie chart.*

**Note**: Make sure exactly one team member submits your in-class activity document [here](https://forms.gle/TdJ3m38QVYSkXWpX8) *before* you leave lecture today!

**Exercise 3: Reflect on the Consequences (20 mins)**

In this exercise, we will reflect on how the decisions we made in our visualizations (both earnest and deceptive) may help or harm our intended visualization **audience**, as well as the subjects being **analyzed**. Please use the [ethical considerations “cheat sheet”](https://docs.google.com/presentation/d/1oEHt1hi9Do8NxC2EAUn7uf9H-cl37L98t2qP1OTMYGU/view) and to guide your team’s discussion. According to these slides, we are the “analysts” creating visualizations about certain “subjects” (here, workers in the US) for a target “audience” (let’s say the general public).

Take notes on your group’s discussion below:

*This visualization suggests that because the men and women pie charts look similar, which falsely implies there is no gender inequality in earnings. Voices left out are nonbinary folks, and folks with races not represented in the data such as native american, etc (as well as multiple races). On the visualization side, intentional because pie charts normalize to 100%. On the data collection side, the data collectors choose who to sample and how to represent them.*

**Note**: Make sure exactly one team member submits your in-class activity document [here](https://forms.gle/TdJ3m38QVYSkXWpX8) *before* you leave lecture today!

**Exercise 4: Final Reflection (5 min)**

What are some takeaways that you took from these exercises today? How might you use what you learned to complete A2?

* *The same data can tell a different (and misleading) story, depends on the intent of the creator*
* *Comparative pie charts are often misleading, hard to compare*
* *Normalizing data can often be deceptive, important to understand what the normalization transformation does and what data is lost*

**Note**: Make sure exactly one team member submits your in-class activity document [here](https://forms.gle/TdJ3m38QVYSkXWpX8) *before* you leave lecture today!