Module 5 Visualization Lab

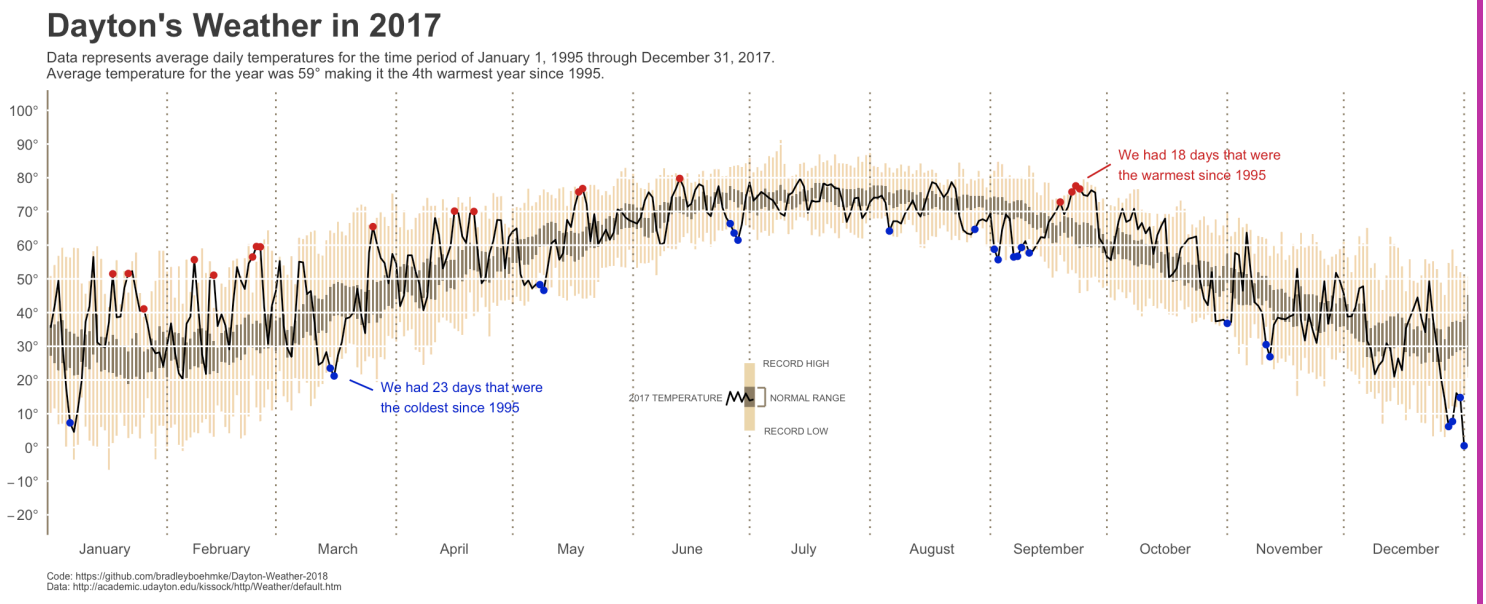
Start Assignment

- Due Nov 12, 2023 by 11:59pm
- Points 10
- Submitting a file upload

Objective

For this module you will use the **completejourney** data (we discussed and used this data in **lesson 4b (https://bradleyboehmke.github.io/uc-bana-6043/_build/html/04-module/lesson-4b.html#prerequisites)**). We'll use the Python **completejourney_py** package; however, you can find comprehensive documentation for this data at **this sister R package (https://bradleyboehmke.github.io/completejourney/articles/completejourney.html)**. The objective is to uncover unique insights in this data that you can answer with exploratory data visualization. You need to develop three informative and professionally formatted plots that align to the grading rubric below and tells a story about the **completejourney** data that we haven't seen yet.

The plots need to tell the entire story, which means you cannot have additional text before or after the image in your Jupyter Notebook. Consequently, you need to use titles, subtitles, captions, x & y axis labels, etc. to appropriately tell the story the image is trying to relay. The image that follows is an example of an image that tells a story without requiring additional text outside the image.



This is your opportunity to be creative and to dig into this data set while building your visualization skills!

- Contains three distinct plots (**2pts**)
- Code for each plot must be visible and follows proper coding style. (**1pt**)
- You use 2 or more of the **completejourney** data sets for each plot. This means you are required to do some data joining for each plot. (**1pt**)
- You must do some filtering, aggregation, etc. of the data for the plotting. What this means is you can't
  just join transactions with demographics and then plot sales_value vs. quantity for each income level. A more interesting find would be to plot the total sales_value per quantity for the top 5 products that each income level tends to purchase. (**1pt**)
- Your plots combine both numeric and categorical features. (**1pt**)
- Each plot explains the story it is trying to tell with and only with plot features (i.e. title, subtitle, x & y axes, annotation, etc.) and no additional prose in the Jupyter Notebook. (**1pt**)
- Each plot is visually appealing, simple to understand, and not overly cluttered. (**1pt**)
- Each plot is appropriately sized meaning the dimensions appear appropriate and text (i.e. titles, axis labels & tick marks) are not overlapping. (**1pt**)
- The story the plot is telling is creative and has not been uncovered in other class sessions/exercises. (**1pt**)

Submission

This module's lab will be a little different than the previous modules. You will be uploading your rendered Jupyter Notebook HTML output with the plots to this assignment location. Be sure to have you name in the Notebook title (i.e. week5_lab_bcboehmke.html) and that what you are submitting is the rendered notebook HTML output and not the .ipynb file that needs to be executed. If you submit the .ipynb file that needs to be executed and your reviewers can't execute it then you will get zero points for completion.

Peer review

This lab will be peer reviewed for the grading. This means 5 of your classmates will be reviewing your plots and using the above rubric to determine your score. Consequently, your objective is to not only impress me but also your classmates with your visualizations.

This also means you will be selected to grade several of your peers submissions. **If you fail to submit your reviews you will automatically be deducted 25% on your overall engagement portion of the class grade.**

To complete the peer review, **follow these instructions (https://community.canvaslms.com/t5/Student-Guide/How-do-I-submit-a-peer-review-to-an-assignment/ta-p/293)** . When completing the peer review be sure to fill out the grading rubric (as described in the

Module 5 Visualization Lab

Criteria	Ratings		Pts
Contains three distinct plots	**2 pts Full Marks**	**0 pts No Marks**	2 pts
Code for each plot must be visible and follows proper coding style	**1 pts Full Marks**	**0 pts No Marks**	1 pts
You use 2 or more of the completejourney data sets for each plot. This means you are required to do some data joining for each plot.	**1 pts Full Marks**	**0 pts No Marks**	1 pts
You must do some filtering, aggregation, etc. of the data for the plotting. What this means is you can't just join transactions with demographics and then plot sales_value vs. quantity for each income level. A more interesting find would be to plot the total sales_value per quantity for the top 5 products that each income level tends to purchase	**1 pts Full Marks**	**0 pts No Marks**	1 pts
Your plots combine both numeric and categorical features	**1 pts Full Marks**	**0 pts No Marks**	1 pts
Each plot explains the story it is trying to tell with and only with plot features (i.e. title, subtitle, x & y axes, annotation, etc.) and no additional prose in the Jupyter Notebook	**1 pts Full Marks**	**0 pts No Marks**	1 pts
Each plot is visually appealing, simple to understand, and not overly cluttere	**1 pts Full Marks**	**0 pts No Marks**	1 pts
Each plot is appropriately sized meaning the dimensions appear appropriate and text (i.e. titles, axis labels & tick marks) are not overlapping	**1 pts Full Marks**	**0 pts No Marks**	1 pts

Criteria	Ratings		Pts
The story the plot is telling is creative and has not been uncovered in other class sessions/exercises	**1 pts Full Marks**	**0 pts No Marks**	1 pts
			Total Points: 10