COMP 4304 / 6934 – Assignment 5 (6%)

MEMORIAL UNIVERSITY

Due: 11:59pm, Mar. 4, 2024

Learning Objectives

The goal of this assignment is to become familiar with radial designs.

Data

Download the `births.csv` and `consoles.csv` data sets from Brightspace.

The `consoles.csv` data set contains information on video game consoles released by Nintendo, Microsoft and Sony over the past few decades. It has the release date for each console, the type of console (home, handheld, etc), and the number of units that were sold.

The `births.csv` data set contains information on the date and time of every birth in the USA in the year 2022. This data set originates from the National Center for Health Statistics, which is part of the Center for Disease Control (https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm).

You can read more about the original visualization based on similar data in the following articles:

- Monday, 8 A.M.: Time to Have a Baby, *Scientific American*, https://www.scientificamerican.com/article/monday-8-a-m-time-to-have-a-baby/
- Why Are so Many Babies Born Around 8 A.M.?, *Zan Armstrong*, https://blogs.scientificamerican.com/sa-visual/why-are-so-many-babies-born-around-8-00-a-m/
- The Baby Spike, *Nadieh Bremer*, https://www.visualcinnamon.com/portfolio/baby-spike/.
- Creating the Scientific American Baby Spike Visual, *Nadieh Bremer*, https://www.visualcinnamon.com/2017/10/creating-baby-births-visual/

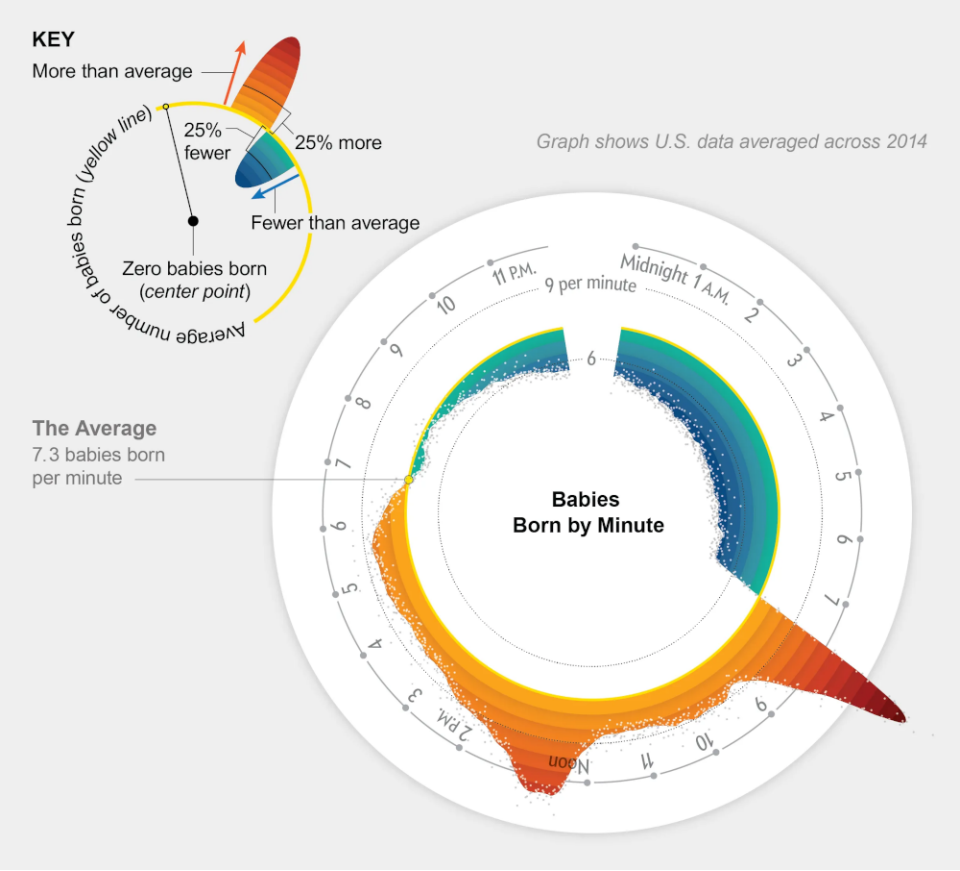
Instructions

Using the provided data sets, create a Jupyter notebook to answer the following questions.

You may only import the pandas, Matplotlib, Seaborn, math or NumPy packages.

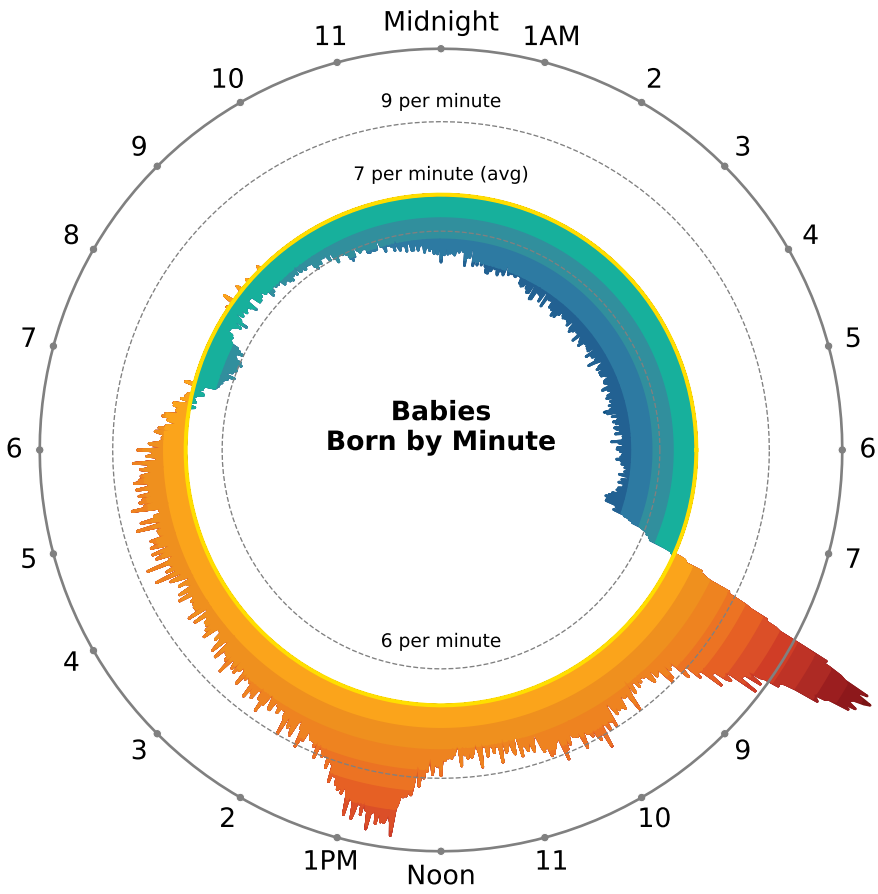
Question 1: (75 pts)

Radial designs often pair well with periodic data. The visualization below from Nadieh Bremer and Zan Armstrong makes use of the periodicity of 24-hour periods to its advantage. It shows the average number of babies born per minute in the USA for the year 2014. There is a large spike of births just after 8am.



The yellow line represents the average number of babies born per minute. There is a gradient colouring from the average to the actual number of babies born per minute, with a blue gradient for times when the actual number is below the average, and a red gradient when the actual number is higher. The gradients use a discrete number of colours, with each three steps corresponding to a 25% increase or decrease from the average (see the legend).

The following visualization is an emulation of this visualization in Matplotlib using 2022 data.



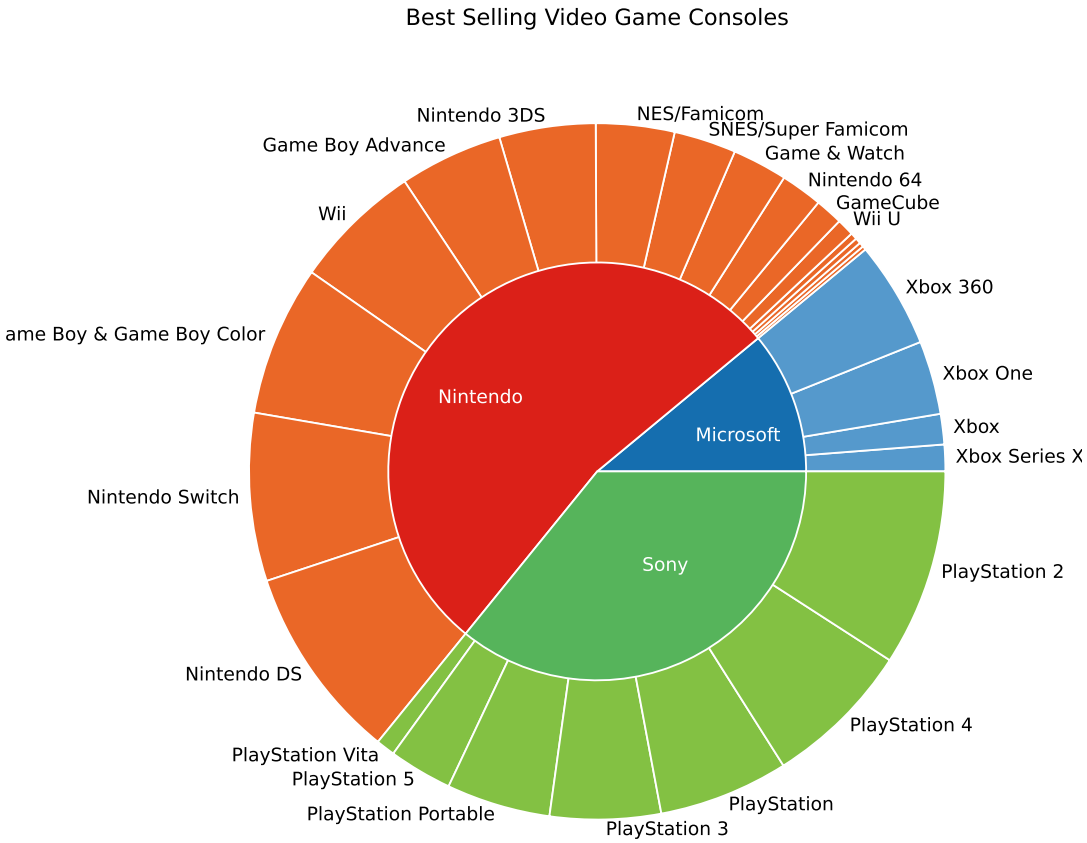
Re-create the emulated visualization using the `births.csv` data set. Several important elements that should be present in your re-creation:

- The region between the actual data and the average line is coloured (hint: fill_between()). Each 3 steps correspond to a 25% increase or decrease from the average.
- The average line is present (colour '#ffde00'), as are the dashed lines for 6 and 9 minutes.
- The angular axis is located at a radius of 11 babies per minute (it intersects the peak, as in the original image). There are circles along the axis for each hour.
- All labels and annotations are present as in the original image.

The gradient colours are ['#17b09c', '#318f9d', '#2d7aa2' '#226192', '#094e81'] for the gradient below the average, and ['#fba41b', '#ef901e', '#ee8320', '#ec7323', '#e66024', '#db4f28', '#d03d26', '#be3326', '#ad2a24', '#9b1e1f', '#8d181b', '#7b1515'] for above the average.

6934 Students Only – Question 2 - Sunburst Chart: (25 pts)

The figure below is a sunburst plot showing the total number of video game consoles sold by Nintendo, Microsoft and Sony. The inner ring shows the total number of consoles sold per company, and the outer ring refines each company into their individual consoles.



Recreate the above sunburst plot using Matplotlib. The arc length of each wedge should be proportional to the number of units sold. Importantly, the outer ring should be aligned with the inner ring. For example, all of the Nintendo consoles on the outer ring should be within the arc spanned by Nintendo.

The consoles on the outer ring should be ordered from largest number of units sold to smallest number of units sold in a clockwise fashion. The four consoles with the smallest number of units sold for Nintendo are not displayed (since that would create visual clutter).

The colours of the inner ring are [#156eaf, #db2018, #56b45b] for Microsoft, Nintendo and Sony, and [#5599cc, #ea6727, #83c143] for the outer ring.

The edge colours of the wedges should be white. The labels of the inner ring should be white in colour and on the interior, and the labels of the outer ring on the exterior.

Submission

Submit your Jupyter notebook (`.ipynb`) through Brightspace. The modified alluvial diagram source code should be embedded inside your notebook (preferably not as a standalone piece of code).

Late submissions will be subject to a 10% penalty for each hour past the deadline.

Attribution

Submissions should include an attribution section indicating any sources of material, ideas or contribution of others to the submission.

Submissions must represent your independent work.

You are encouraged to use any resources to help with your solution, but your solution must represent independent work. If your submitted work includes unacknowledged collaboration, code materials, ideas or other elements that are not your original work, it may be considered plagiarism or some other form of cheating under MUN general regulations 6.12.4.2 (4.12.4.2 for graduate students) and academic penalties will be applied accordingly.

Avoid academic penalties by properly attributing any contribution to your submission by others, including internet sources and classmates. This will also help distinguish what elements of the submission are original. You may not receive full credit if your original elements are insufficient, but you can avoid penalties for plagiarism or copying if you acknowledge your sources.

Github

I encourage you to store and version your work on GitHub. It is good practice to do so as everyone uses git in the real world.

However, **it is a requirement that git repositories containing assignment material be private.** University regulations (undergraduate 6.12.4.2 and graduate 4.12.4.2) consider it cheating if you allow your work to be copied. There will be zero tolerance for this.