

Week 10 Tutorial Worksheet

AY23/24 Semester 2

No submission required

Question 1 Olympics data

In this question, we will work with data on the Olympics from 1896 to 2016 from a TidyTuesday project. You can read the [data description here](#).

1. Read in `olympics.csv` from the TidyTuesday repository as `olympics`. Create a data frame named `summer_games` that contains the number of nations (represented by `noc` in the data) and the number of events per summer game.

The columns of `summer_games` should follow this structure.

```
str(summer_games)
```

```
## tibble [29 x 3] (S3: tbl_df/tbl/data.frame)
##  $ year      : num [1:29] 1896 1900 1904 1906 1908 ...
##  $ n_nations: int [1:29] 12 31 15 21 22 29 29 45 46 47 ...
##  $ n_events  : int [1:29] 43 90 95 74 109 107 158 131 122 131 ...
```

2. Typically, the Olympics are held every four years. Explore the data set and find out the shortest and longest gaps between consecutive summer Olympics.
 - The shortest gap between summer Olympics is _____ years.
 - The longest gap between summer Olympics is _____ years.
3. For each country, compute the number of medals received (gold, silver, bronze, and the total) in each summer game since 1980.

Sort the table in descending order of **golds received**. If there is a tie in gold medals, sort by the number of silver medals. If there is still a tie, break it with the number of bronze medals. Name the final table as `medals`.

The first ten rows of the table should look like the following.

```
head(medals, 10)
```

```
## # A tibble: 10 x 6
##   Country Year Gold Silver Bronze Total
##   <chr>   <dbl> <int>  <int>  <int> <int>
```

##	1	USA	1984	82	61	30	173
##	2	URS	1980	80	69	46	195
##	3	URS	1988	54	31	46	131
##	4	CHN	2008	51	21	28	100
##	5	GDR	1980	47	37	42	126
##	6	USA	2016	46	37	38	121
##	7	USA	2012	46	28	29	103
##	8	EUN	1992	45	38	29	112
##	9	USA	1996	44	32	25	101
##	10	CHN	2012	38	27	23	88

Question 2

During lecture, we covered some principles behind visualization of data. In this question, your task is to **choose ONE** of the graphs below. Comment on its quality/potential problems using the ideas and terminologies covered in class to discuss:

- What does the visualization describe?

Begin by describing what the chosen visualization represents. What data is being conveyed by the graph, and what is the intended message or information?

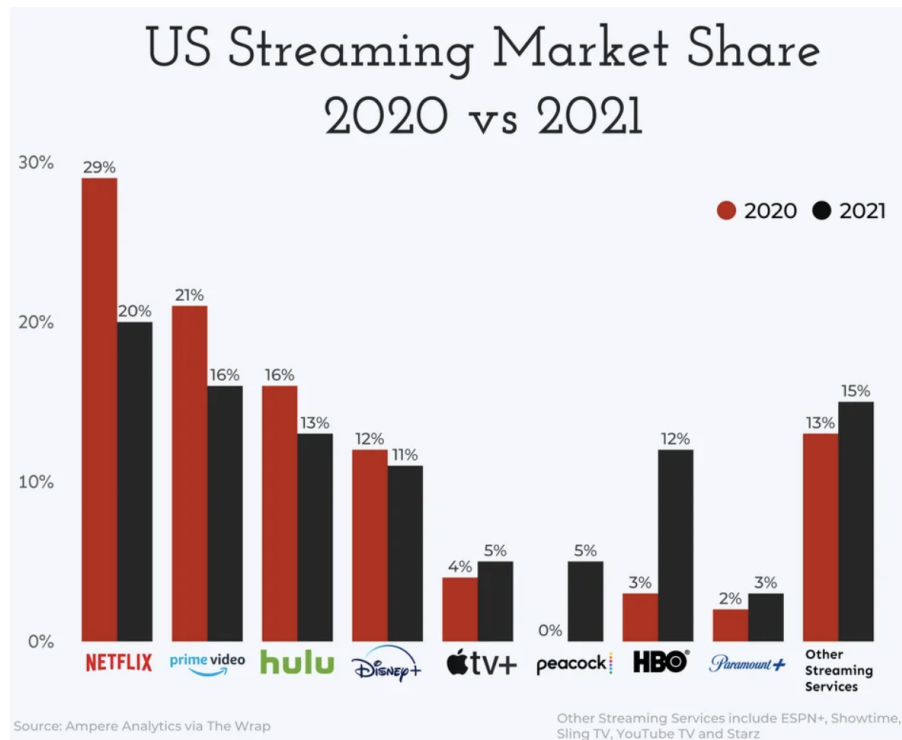
- What do you like/not like about it?

Share your opinions on what you like and dislike about the visualization. This could include aspects like the choice of colors, labels, clarity, and any other visual elements.

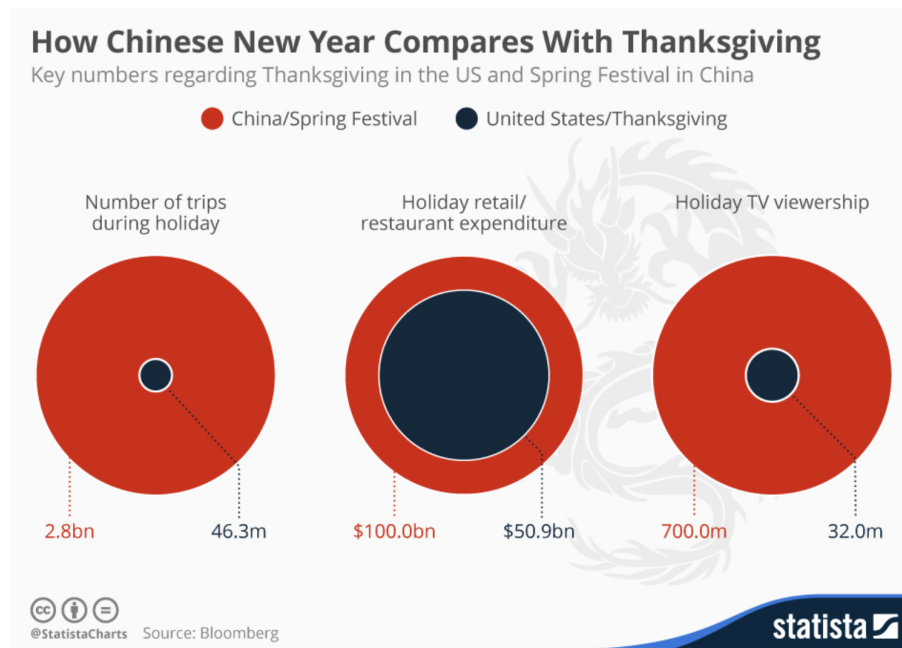
The goal of this question is for you to be able to discuss visualizations intelligently and evaluate others' graphs. Most of the charts are self-explanatory. Comment on the chart based on presented information only. If you find that something is not clear, include it in your comments.

There is no R code needed for this question.

1. Top streaming services in the US, from Reddit.



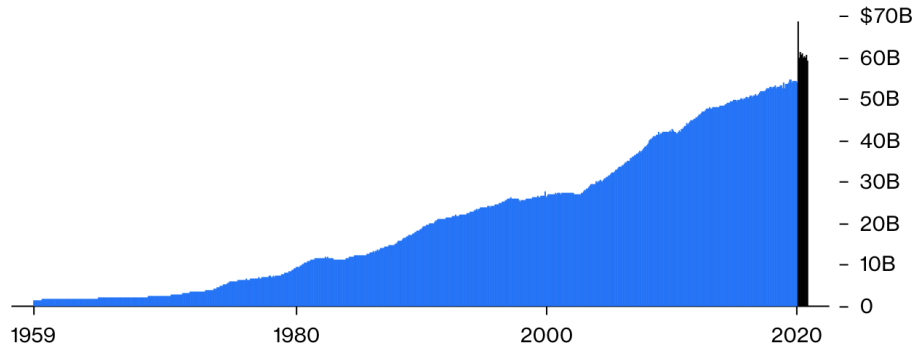
2. How Chinese New Year compares with Thanksgiving, from Bloomberg via Statista.



3. Monthly cereal consumption, covered in Bloomberg.

Well That Sure Was Unusual

Monthly U.S. personal consumption expenditures on cereals*



Source: U.S. Bureau of Economic Analysis

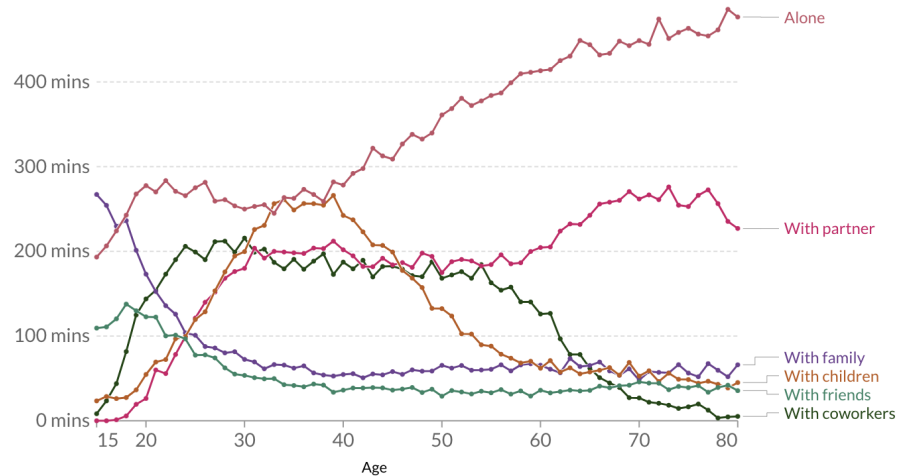
*Includes flour, rice, pasta etc. as well as breakfast cereal, but does not include restaurant meals

4. Who do we spend time with in our life, from Our World in Data.

Who Americans spend their time with, by age

Average time spent with others is measured in minutes per day, and shown by the age of the respondent. This is based on averages from surveys between 2009 and 2019.

Our World
in Data



Source: American Time Use Survey (2009-2019) and Lindberg (2017)

Note: Relationships used to categorize people are not exhaustive. Additionally, time spent with multiple people can be counted more than once (e.g. attending a party with friends and partner counts for both "friends" and "partner").

OurWorldInData.org/social-connections-and-loneliness • CC BY

Requirements

- The code in your Rmd file should create the following objects: `olympics`, `summer_games`, and `medals`.
- No R code is needed for Question 2.