Data Abstraction and Color

Q1: What type of attribute (categorical, ordinal, quantitative) are the following? If the attribute is ordinal or quantitative, what is the ordering direction (sequential, diverging, cyclic)? What color map (#1~#6) is the best choice to encode each attribute?   (30%)

1. Academic letter grade (e.g. A+, A, A-, B+ B ⋯)
2. Media companies (e.g. Netflix, Walt Disney, Fox, CBS, ⋯)
3. Acceleration (e.g. +2 m/s^2 , -3 m/s^2, +22 m/s^2)
4. Scores of an exam (e.g. 90 points, 60 points, 30 points)
5. Survey options (e.g., strongly agree, agree, disagree, strongly, disagree)

Q2: Check this website about mental health survey: (36%)
https://www.kaggle.com/datasets/mahirahmzh/starbucks-customer-retention-malaysia-survey

- Write down the attribute types of the following attributes (categorical, ordinal, quantitative) and explain your choice for every attributes.
    - Id
    - Gender
    - Age
    - Status
    - Income
    - visitNo
    - method
    - timeSpend
    - location
    - membershipCard

- spendPurchase
- produceRate

Visual Channel and Mark:

Q3: You will identify data items and the marks used to encode them, and data attributes and the channels used to encode them. (34%)
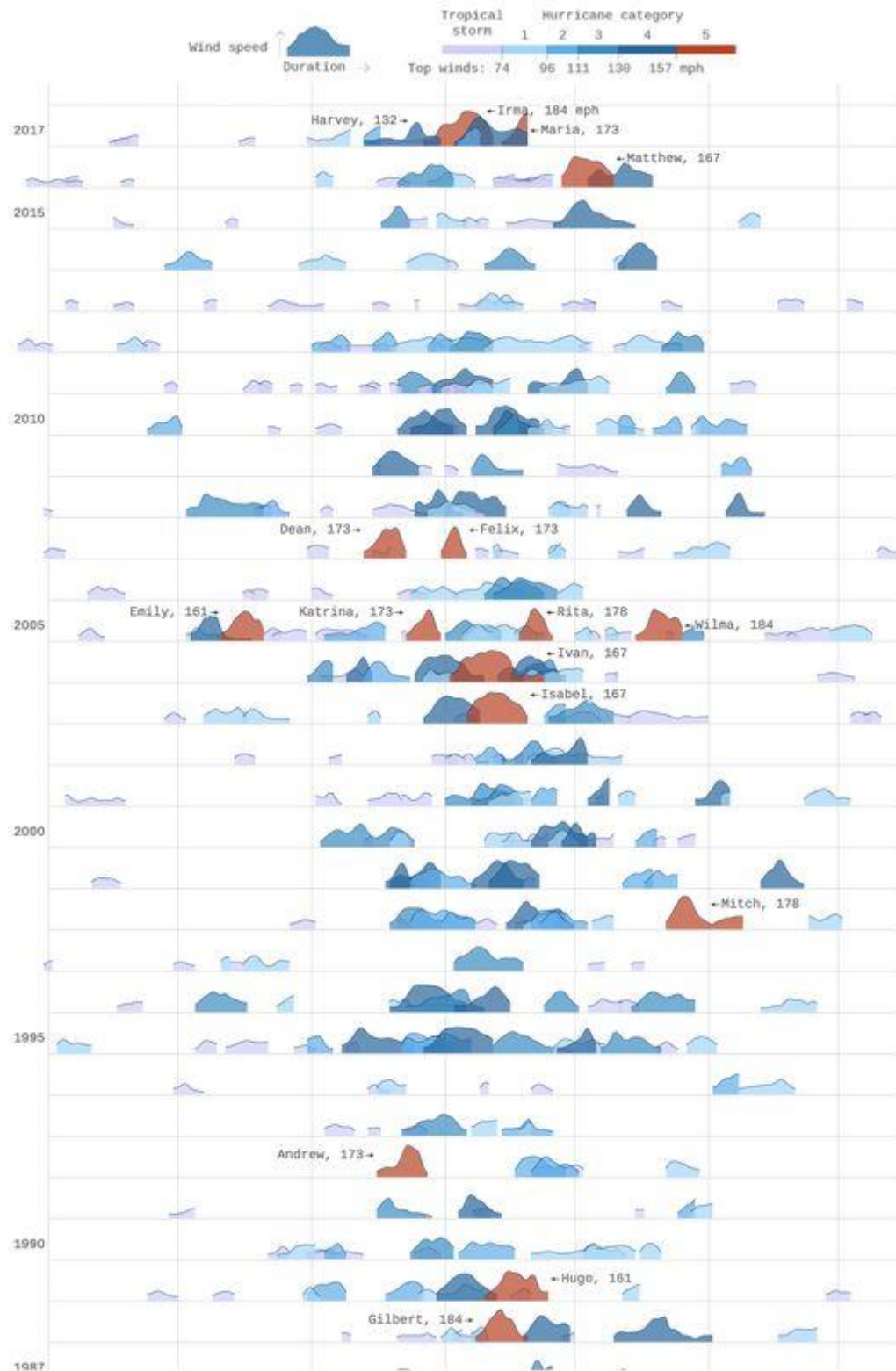
For each chart, fill in
- Visual channels used?
  - Channel X encodes attribute Y
  - Channel X encodes attribute Y
  - ....
- Marks used?
  - Mark of type X encodes item Y
  - Mark of type X encodes item Y
  - …

Note that underneath each chart there is a link to a web page providing details about the project, including information about the data and in many cases supporting interactivity. For now, just encode what you see in the **static image** in this document; you might notice that interactivity is being used a lot for emphasis (hover highlight, animated transitions, reordering, details popups, small multiples zooming). Think about how you would reconstruct the dataset underlying this picture. Assume one mark per item, think of items as rows of a table (or nodes in a network). Assume each channel is showing an attribute, where channel is a column in a simple data table or network. Remember that multiple channels could redundantly encode the same attribute.

A.

Thirty years of Atlantic hurricanes

This is a history of every Atlantic storm tracked by the National Oceanic and Atmospheric Administration since 1987.

B: https://archive.nytimes.com/www.nytimes.com/interactive/2013/05/25/sunday-review/corporate-taxes.html

C: https://projects.fivethirtyeight.com/2020-election-forecast/

Quiz credit: Tamara Munzner