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? (20%)

1. Academic letter grade (e.g. A+, A, A-, B+ B …)

**A: Ordinal, ordered with sequential direction, best using #3.**

1. Media companies (e.g. Netflix, Walt Disney, Fox, CBS, …)

**A: Categorical, best using #4.**

1. Acceleration (e.g. +2 m/s2 , -3 m/s2 , +22 m/s2 )

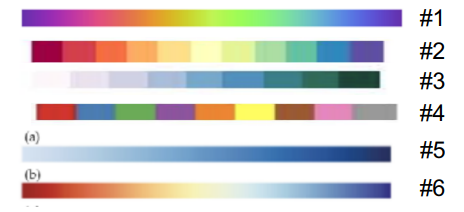
**A: Quantitative, ordered with diverging direction, best using #6.**

1. Scores of an exam (e.g. 90 points, 60 points, 30 points)

**A: Quantitative, ordered with sequential direction, best using #5.**

1. Survey options (e.g., strongly agree, agree, disagree, strongly, disagree )

**A: Ordinal, ordered with diverging direction, best using #6.**



Q2: Check this website about mental health survey dataset : (20%) https://www.kaggle.com/osmi/mental-health-in-tech-survey

Answer questions for the following attributes(columns), timestamp, age, gender, country, state, work\_interfere, no\_employees, benefits, wellness\_program, comments.

* Write down the attribute type (categorical, ordinal, quantitative)
* Determine its cardinality (number of levels) for categorical or ordinal, or range for quantitative

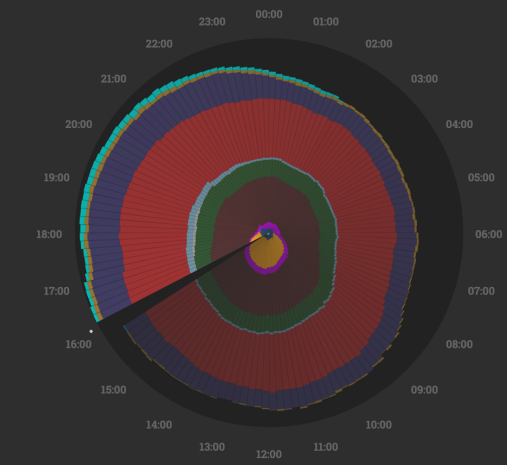
|  |  |  |
| --- | --- | --- |
| Category | Attribute type | Cardinality |
| timestamp | quantitative | 27Aug14 ~ 2Feb16 |
| age | quantitative | -1726 ~ 99999999999  (we can see that there are lots of data that make nonsense here…) |
| gender | categorical | 49 levels  (Maybe the author didn’t limit the input type, so many different types are actually the same one. e.g. ‘M’, ‘Male’, ‘male’, ‘m’ are all represent “male”) |
| country | categorical | 48 levels |
| state | categorical | 46 levels |
| work\_interfere | ordinal | 5 levels (Never, rarely, sometimes, often, NA.  Since we got never here, NA means I don’t know, I guess) |
| no\_employees | quantitive | 6 levels |
| benefits | ordinal | 3 levels (Yes, no, don’t know) |
| wellness\_program | ordinal | 3 levels |
| comments | categorical | 161 levels |

Q3: You will identify data items and the marks used to encode them, and data attributes and the channels used to encode them. (4 questions) (60%)

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
  + ….

A. https://pm25.lass-net.org/Power/TW/index.html?fbclid=IwAR0cqB6FsjoO8-- XGdgEoeYTfpJ4MmoNnMQ8tM6lAqUcRfslEH6EDX1kpfs



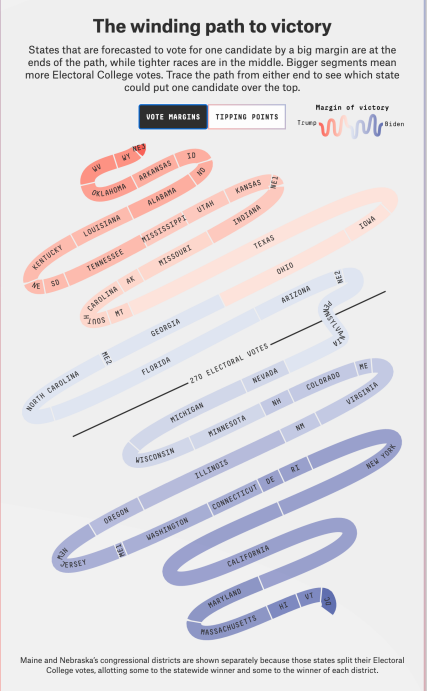
**A: Channels:**

* **Direction (angle to the top) encodes the daytime**
* **Color encodes the types of electric generation in Taiwan**
* **Bar height (length) encodes million watts of electric generated from the type**

**Marks:**

* **Bars encodes million watts the electric generated from that time**

B: <https://projects.fivethirtyeight.com/2020-election-forecast/>

****

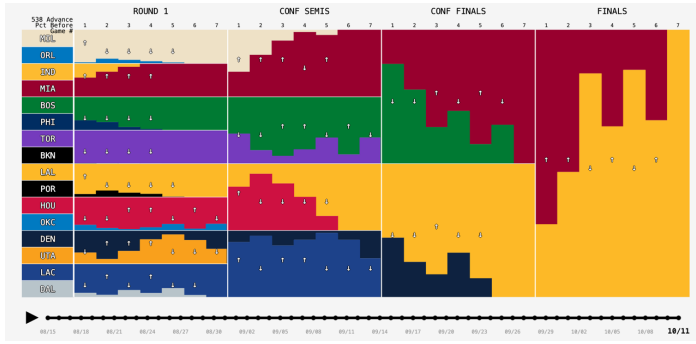
**A: Channels:**

* **Color encodes the percentage of tickets the candidate got.**
* **Length encodes the ticket number the candidate got.**
* **Position in line encodes the percentage of tickets the candidate got.**

**Marks:**

* **Line segment encodes states of the U.S..**

C: <https://roadtolarissa.com/playoff-probabilities/> (the image on 10/11)

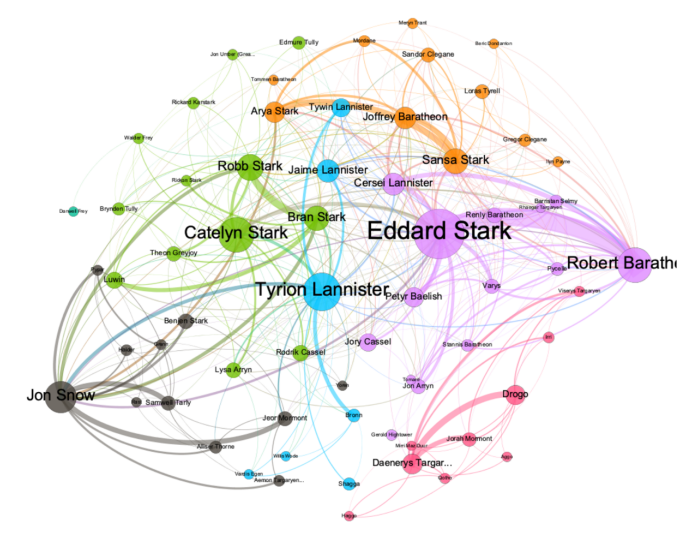
****

**A: Channels:**

* **Color encodes teams.**
* **Row encodes different parts of playoff games.**
* **Column encodes teams.**
* **The bar height encodes the winning probability of the team.**
* **X axis encodes time & games.**

**Marks:**

* **Bars encodes a team’s winning probability that plays in the certain part of the playoff.**
* **Arrows encodes the trend of the team’s predicted winning probability since the last game.**

D: [https://ericmjl.github.io/Network-Analysis-Made-Simple/05-casestudies/01- gameofthrones/](https://ericmjl.github.io/Network-Analysis-Made-Simple/05-casestudies/01-%20gameofthrones/) (you need to read this website to answer this question) 

**A: Channels:**

* **Color encodes seasons of the drama.**
* **Edge weight encodes number of interactions between two characters.**

**Marks:**

* **Points encodes characters.**