DATA 23700 Autumn 2023

Name: Yuyan Wang

CNetID: wangyuyan03

Your CNetID should be the part of your UChicago e-mail address before `@uchicago.edu`.

Exercise 1: Grammar of graphics

Due October 6, 2023

The purpose of this exercise is to gain practice reasoning about visualizations in terms of the grammar of graphics. This entails being able to come up with a formal chart specification for any given chart.

You will consider a series of visualizations and produce specifications for constructing each of them. Since we haven't introduced any particular programming tool for specifying visualizations, we will use psuedocode specifications for this exercise. Your psuedocode specifications should look like this:

```
{
    x: <variable on the x-axis>,
    y: <variable on the y-axis>,
    color: <variable encoded by color>,
    size: <variable encoded by size>,
    angle: <variable encoded by angle>,
    row: <variable on rows of a faceted plot>,
    column: <variable on columns of a faceted plot>,
    layer: [
        {
            mark: <geometry we are using to display data>,
            agg: <aggregation function applied to data>
        },
        ...
        {
            mark: <geometry we are using to display data>,
            agg: <aggregation function applied to data>
        }
    ]

}
```

Notice that there are a variety of encodings that we can use, which apply to all layers of the chart (unless otherwise specified). Also, notice that charts can have one or more layers where different markings are superimposed.

Use this descriptive notation to say how each chart is constructed to the best of your ability. Feel free to supplement the specifications you produce with notes to explain your reasoning if you are struggling with the notation. We are not necessarily looking for perfect specifications,

especially since this is psuedocode. We just want you to practice thinking about chart construction in a formal way.

Note that some of the visualizations provided are missing labels that would tell you which variables are mapped to each encoding channel. In these cases, just indicate which encodings were used to create the visualization, and feel free to leave the variable names blank (e.g., x :???).

Since this is an exercise, you may collaborate with your classmates. However, you must submit your own work to receive credit for the exercise. Feel free to ask your TAs and instructor clarifying questions during the activity!

Please save a local copy of this PDF to do your work in, and to turn in on Gradescope for credit.

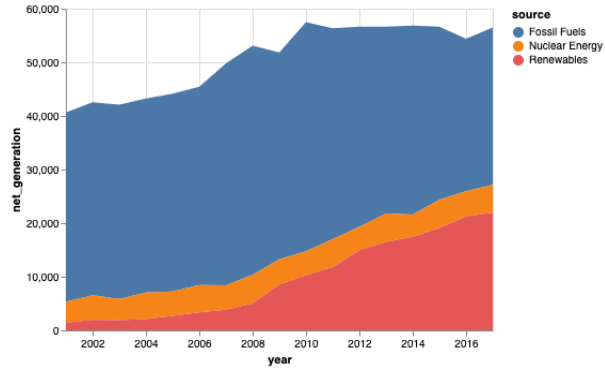---

Problem 1
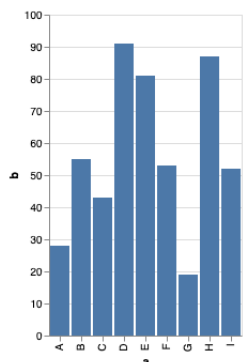

```
{
  x: a (character)
  y: b (int)
  layer: [
    {
      mark: bar
    }
  ]
}
```

Problem 2
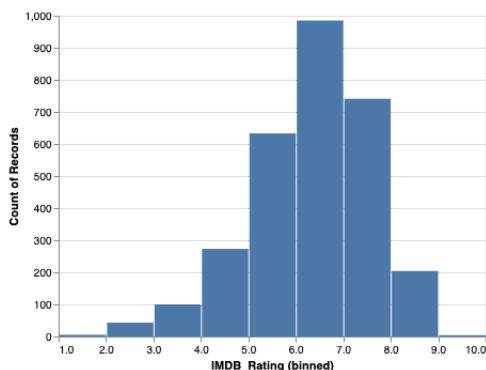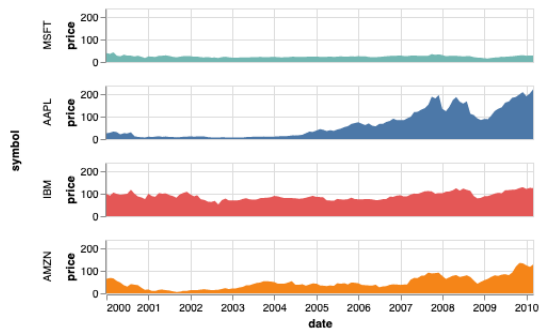


```
{
  x: IMDB_Rating (binned) (floats)
  y:  Count of Records (int)
  layer: [
    {
      mark: bar
      agg: Histogram
    }
  ]
}
```

Problem 3
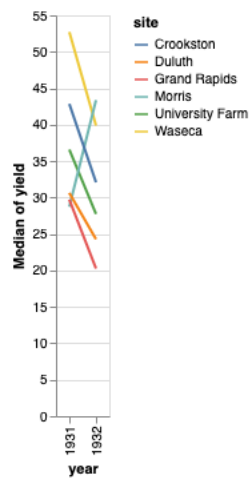


```
{
  x: year
  y: net_generation
  color: source
  layer: [
      {
        mark: area
      }
  ]
}
```

Problem 4
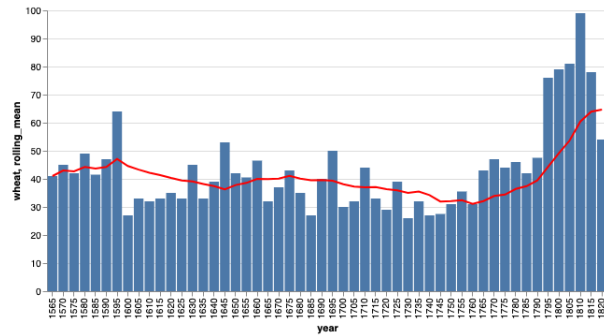


```
{
  x: date
  y: price
  row: symbol
  color: symbol
  layer: [
      {
        mark: area
      }
  ]
}
```

Problem 5


```
{
  x: year
  y: Median of yield
  color: site
  layer: [
      {
        mark: line
      }
  ]
}
```
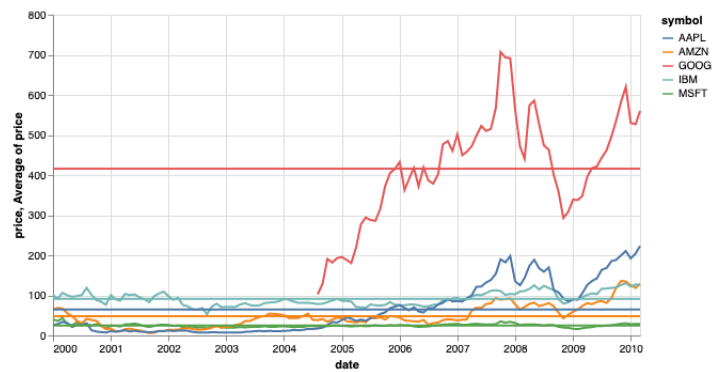
Problem 6
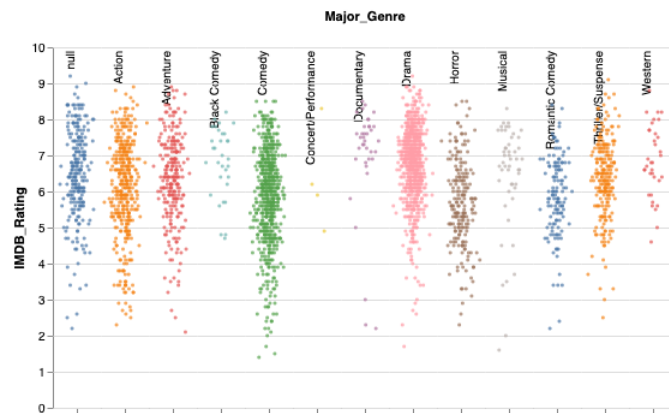


```
{
  x: year
  y: wheat, rolling_mean
  layer: [
    {
      mark: bar
    },
    {
      mark: line
      agg: expectation
    }
  ]
}
```
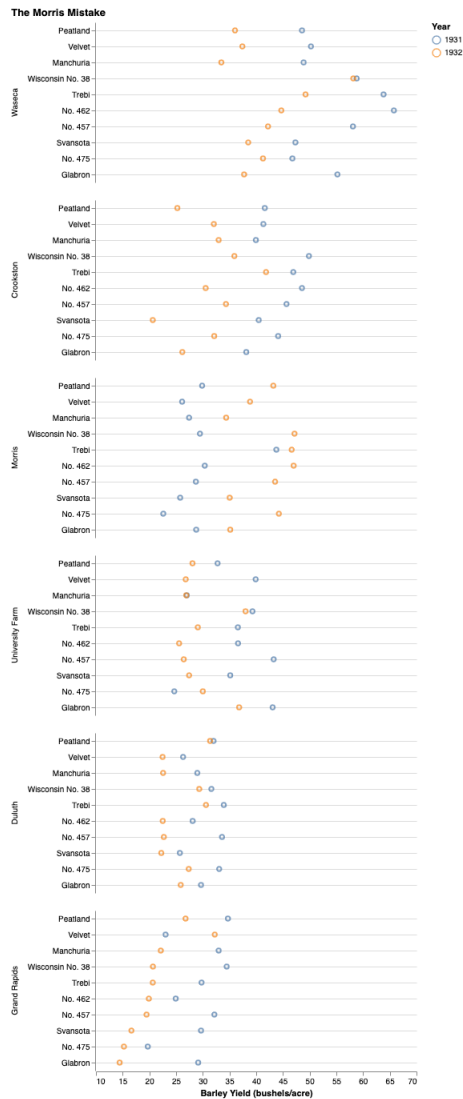
Problem 7


```
{
  x: date
  y: price, Average of price
  color: symbol
  layer: [
    {
      mark: line
    },
    {
      mark: line
      agg: mean
    }
  ]
}
```

Problem 8


```
{
  x: Major_Genre
  y: IMDB_Rating
  color: Major_Genre
  layer: [
    {
      mark: dots
    }
  ]
}
```
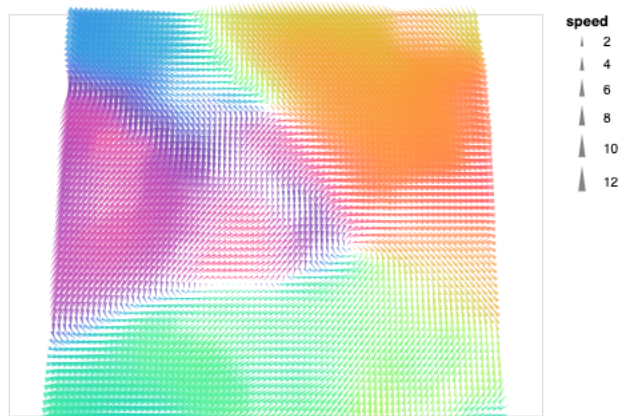
Problem 9


```
{
  x: Barley Yield (bushels/acre)
  y: Variety of barley
  rows: Site in Minnesota
  title: The Morris Mistake
  color: year
  layer: [
    {
      mark: circles
    }
  ]
}
```

Problem 10


```
{
  x: ???
  y: ???
  color: ???
  size: speed
  angle: ??? (guess: direction)
  layer: [
    {
      mark: triangle
    }
  ]
}
```