

Project 3 - 3D Visualization

Jiaheng Tang

Concept Description

For this project, I decided to explore the most popular checked out items of each year, from 2006 to 2018. I picked the top 50 most popular item of each year, as well as the top 50 of the total checkouts.

MySQL Queries

I first constructed my SQL query to get the top 1000 most checkout items, below is the query I used.

```
SELECT COUNT(itemNumber) AS Counts, title, itemNumber, cout, itemType,
deweyClass
FROM spl_2016.inraw
WHERE YEAR(cout) BETWEEN 2006 AND 2018
GROUP BY itemNumber,title, cout, itemType, deweyClass
ORDER BY Counts DESC
LIMIT 1000
```

This query took really long, for about 45 minutes. I then ran a query to get the popular checkout items each year.

```
SELECT COUNT(itemNumber) AS Counts, title, itemNumber, cout, itemType,
deweyClass
FROM spl_2016.inraw
WHERE YEAR(cout)=2018
GROUP BY itemNumber,title, cout, itemType, deweyClass
ORDER BY Counts DESC
LIMIT 50
```

I just changed the `YEAR(cout)` to each year. But I realized that I should get each checkout records of the top 50 most popular items to make it more interesting. Therefore I wrote a Python script to read the item number from the CSV file I got back, and dynamically construct the SQL query I needed. Here is the Python script,

```
import csv
import sys

def main(filename: str):
    with open(filename, 'r') as fin:
        with open(f'cout_{filename[:-4]}.sql', 'w+') as fout:
            year = filename[6:10]
            csv_reader = csv.DictReader(fin, delimiter=',')
            item_num = ''
```

```

for row in csv_reader:
    item_num += f" OR itemNumber={row['itemNumber']}"
    item_num = item_num[4:]
    fout.write(
        f'SELECT title, itemNumber, cout, cin, itemType,
deweyClass\nFROM spl_2016.inraw\nWHERE ({item_num}) AND YEAR(cout)=
{year}\nGROUP BY itemNumber, title, cout, cin, itemType, deweyClass\n')

if __name__ == '__main__':
    main(str(sys.argv[1]))

```

and one of the queries it produced

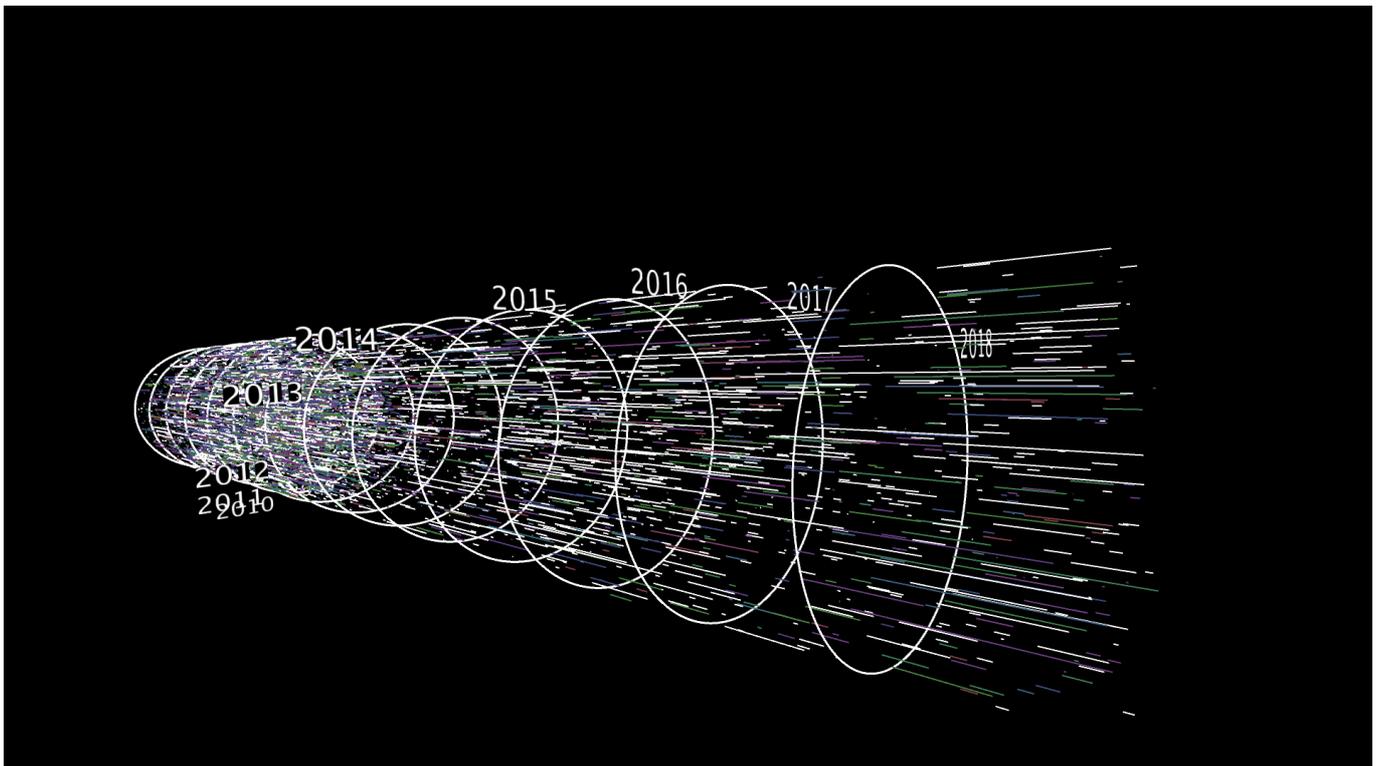
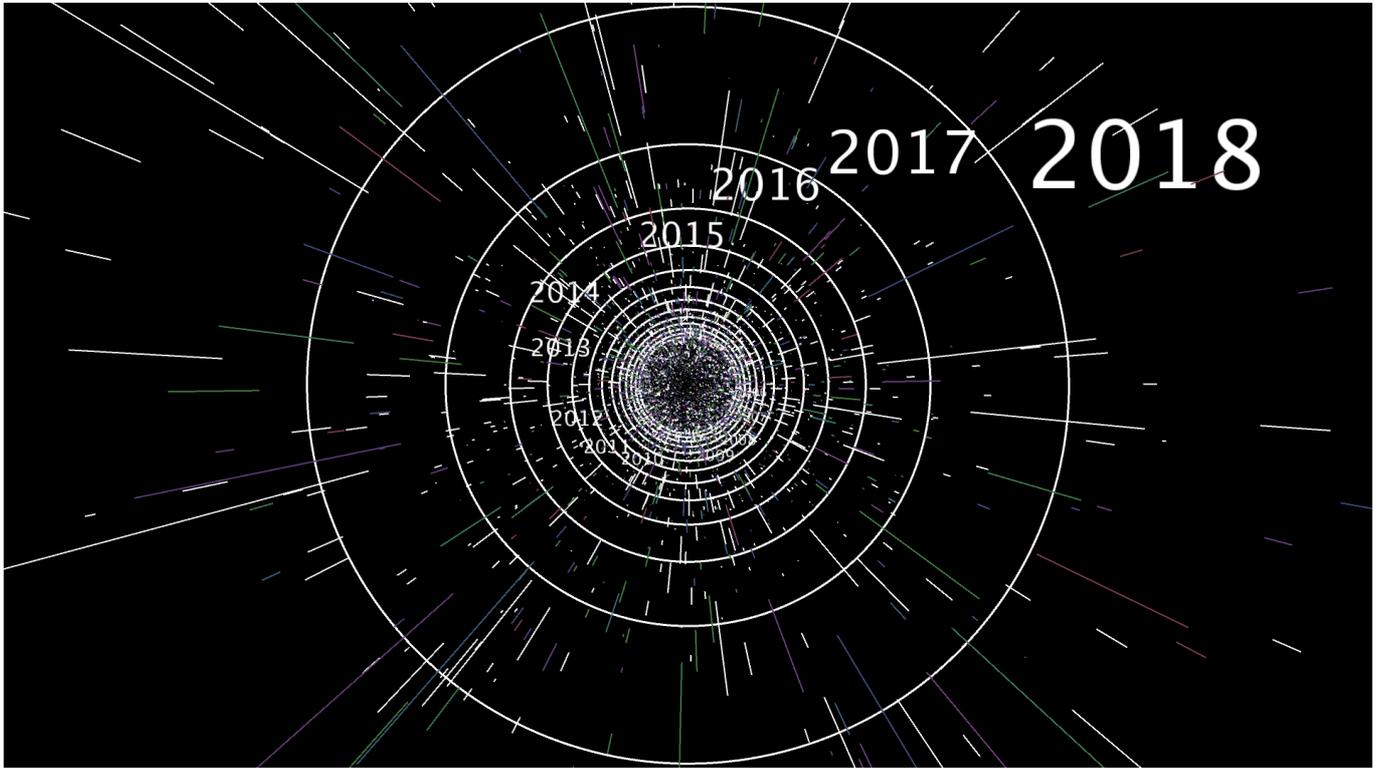
```

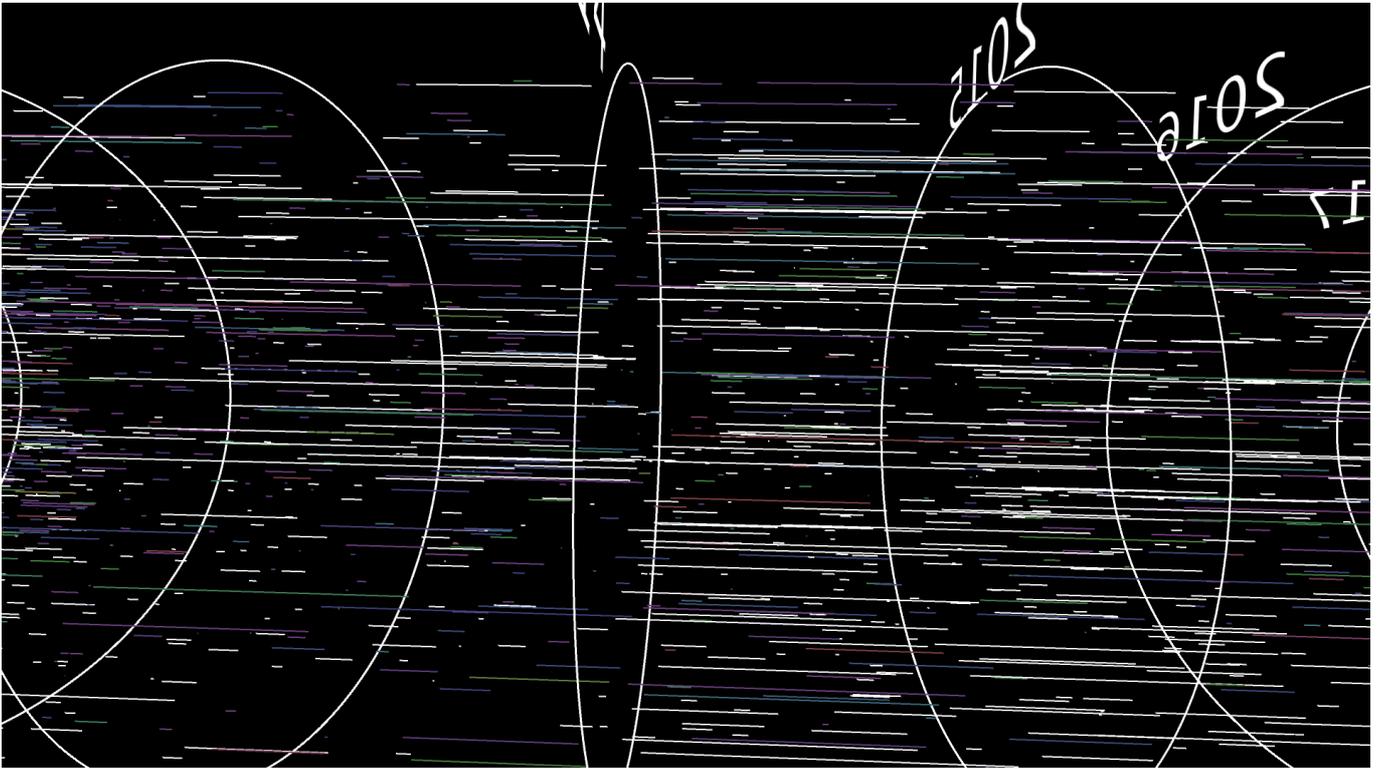
SELECT title, itemNumber, cout, cin, itemType, deweyClass
FROM spl_2016.inraw
WHERE (itemNumber=6168318 OR itemNumber=4814851 OR itemNumber=937827 OR
itemNumber=5446948 OR itemNumber=5456901 OR itemNumber=5299584 OR
itemNumber=3702279 OR itemNumber=3287481 OR itemNumber=4181717 OR
itemNumber=4501511 OR itemNumber=2498093 OR itemNumber=4041489 OR
itemNumber=2566498 OR itemNumber=4795932 OR itemNumber=5524467 OR
itemNumber=4692088 OR itemNumber=4955520 OR itemNumber=5099578 OR
itemNumber=4247243 OR itemNumber=4424525 OR itemNumber=3177807 OR
itemNumber=5317788 OR itemNumber=5248119 OR itemNumber=5082517 OR
itemNumber=5119956 OR itemNumber=2817060 OR itemNumber=5032075 OR
itemNumber=5140938 OR itemNumber=4568479 OR itemNumber=3864614 OR
itemNumber=5464418 OR itemNumber=3760081 OR itemNumber=4910008 OR
itemNumber=4876121 OR itemNumber=2559433 OR itemNumber=5445782 OR
itemNumber=4990761 OR itemNumber=5517891 OR itemNumber=3213417 OR
itemNumber=4441172 OR itemNumber=3572168 OR itemNumber=3494842 OR
itemNumber=4281289 OR itemNumber=1360002 OR itemNumber=4784028 OR
itemNumber=4758313 OR itemNumber=3601773 OR itemNumber=5505648 OR
itemNumber=5007385 OR itemNumber=5285182) AND YEAR(cout)=2018
GROUP BY itemNumber, title, cout, cin, itemType, deweyClass

```

Design

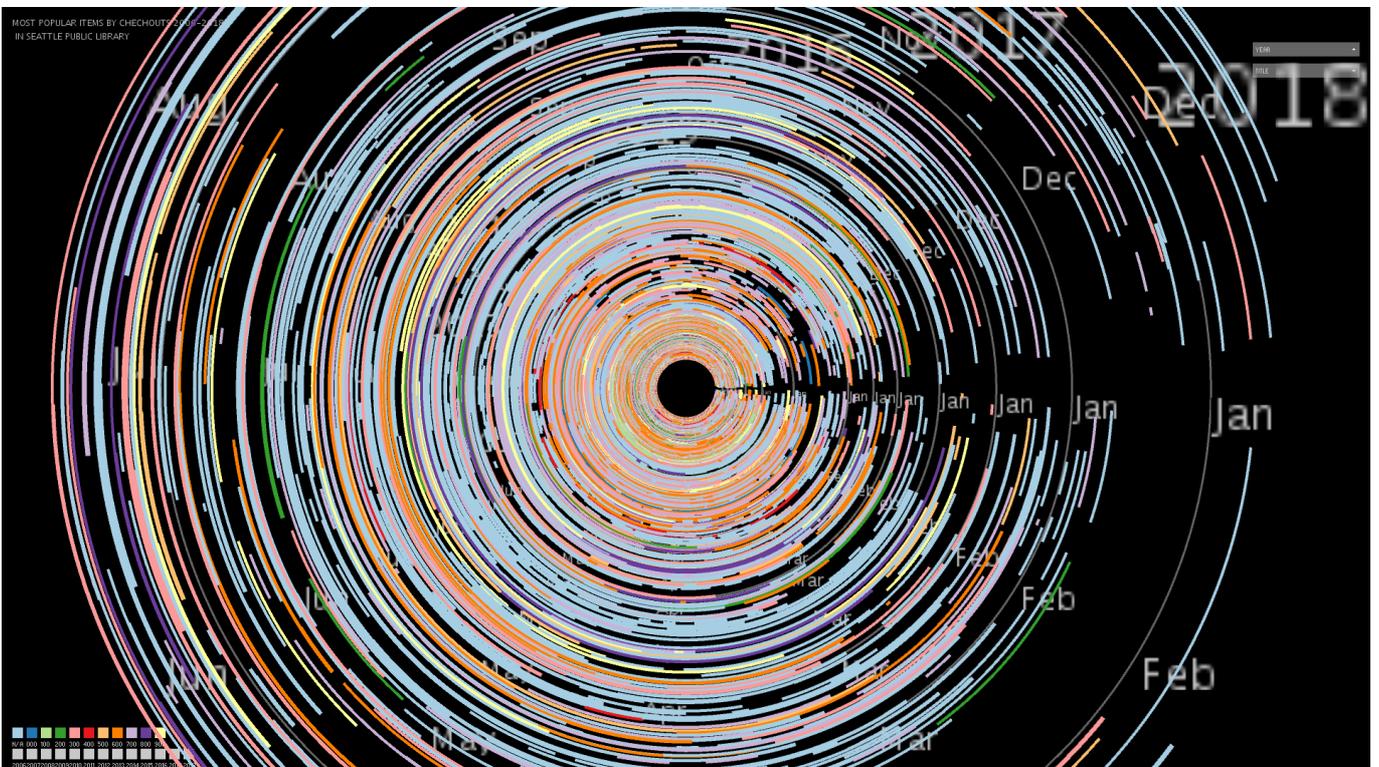
I started out with the idea that I want to use a tunnel like shape to represent time passing by. Then I will plot each record as a line, from the checkout date to check-in date. Here's my first draft.

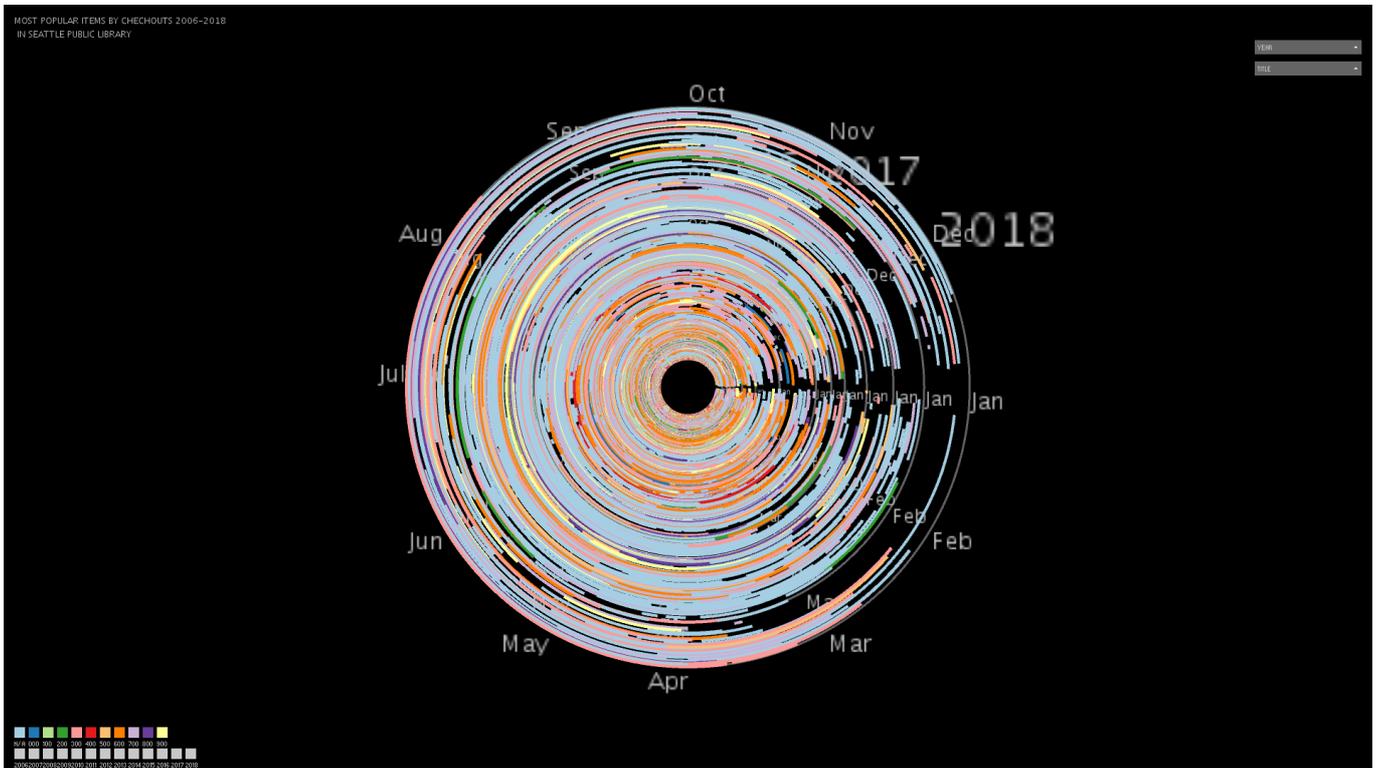


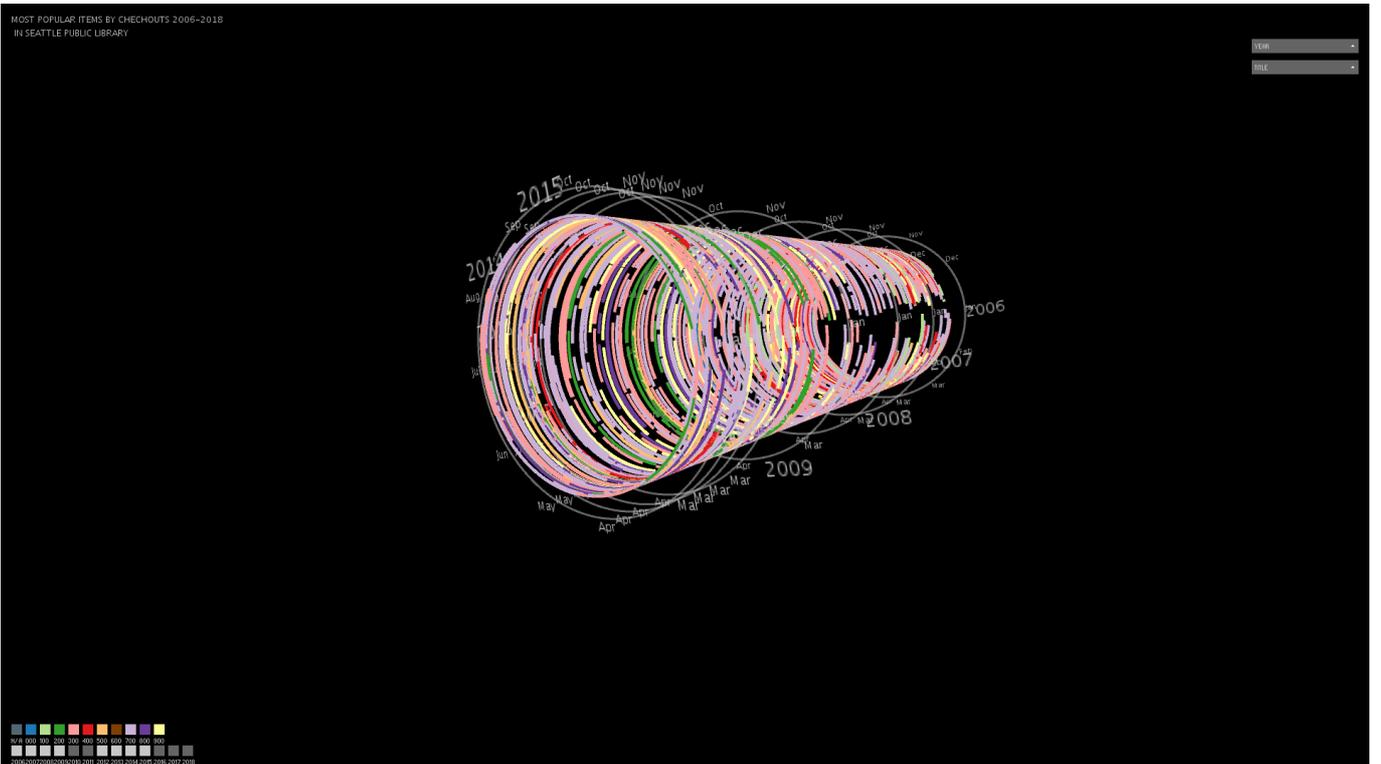
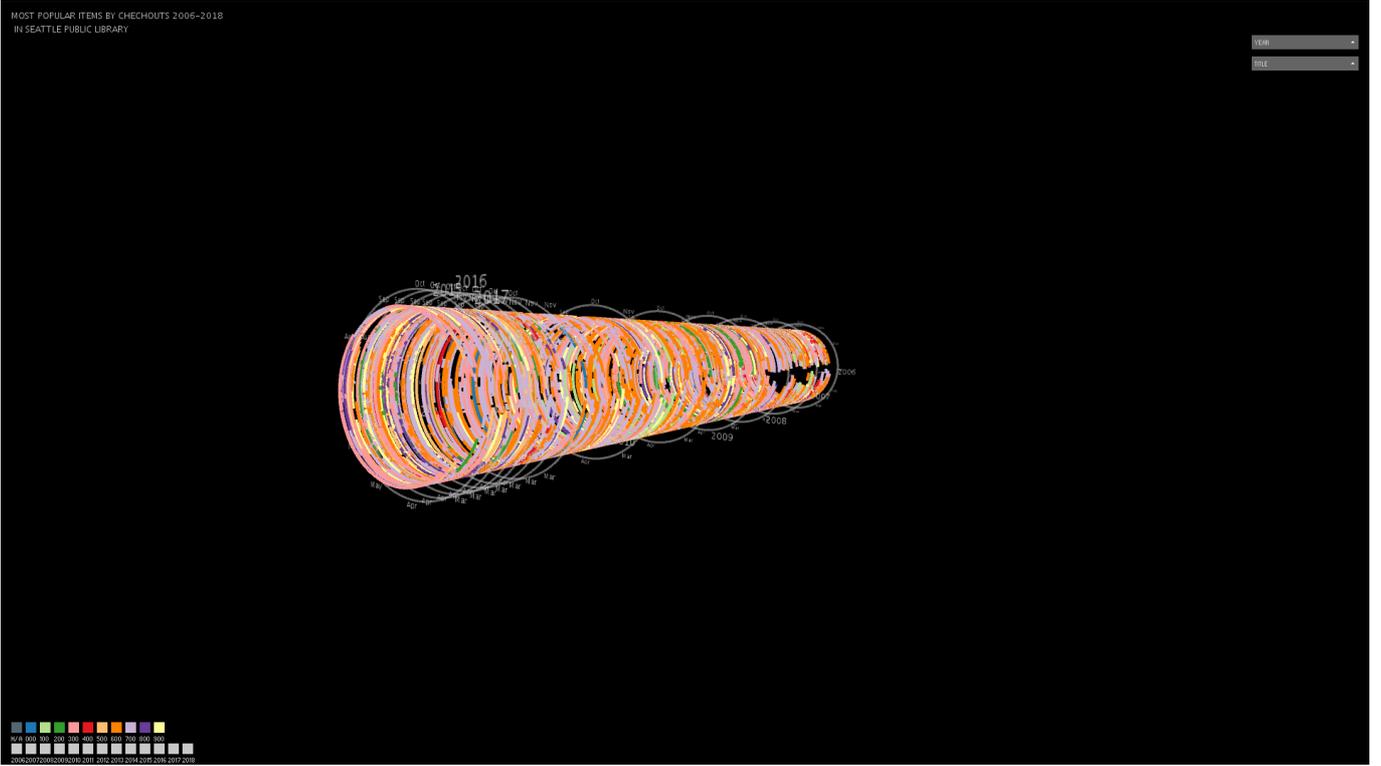


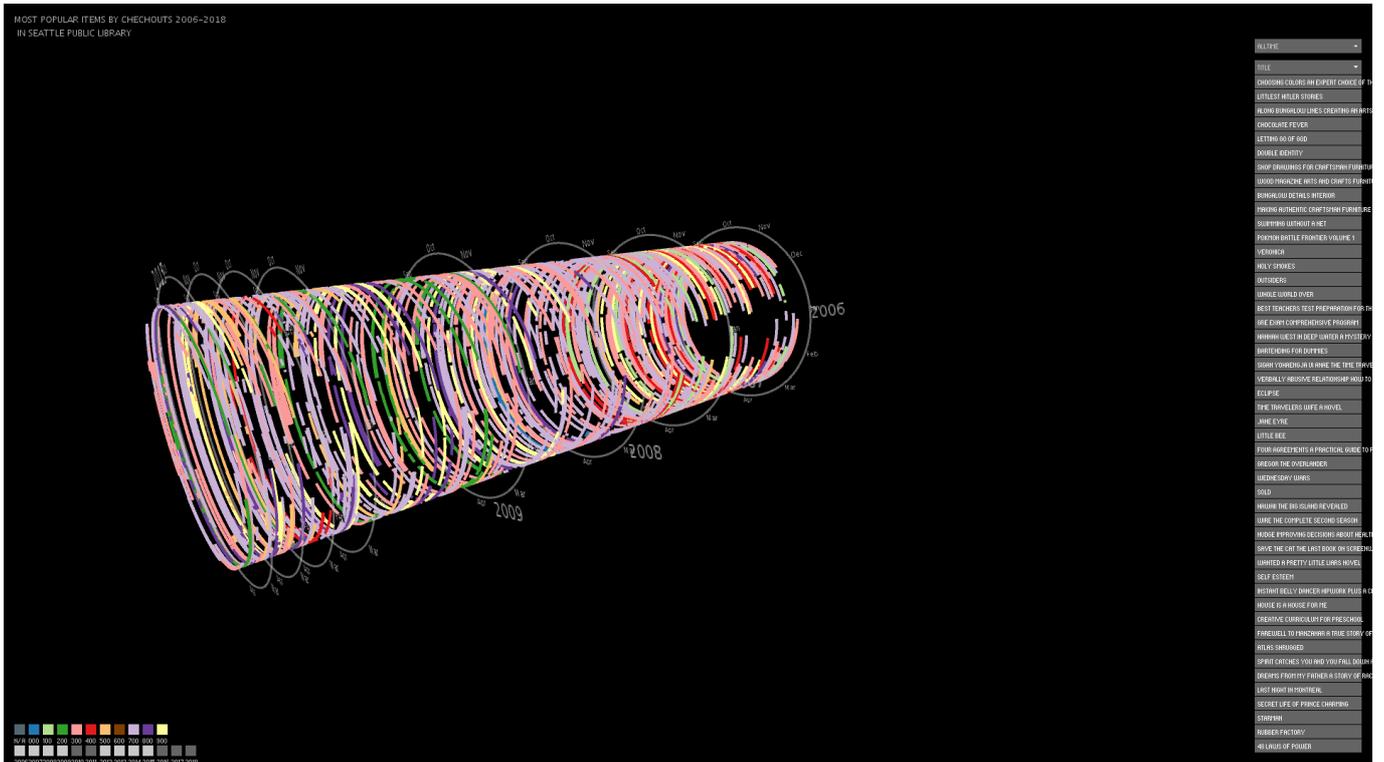
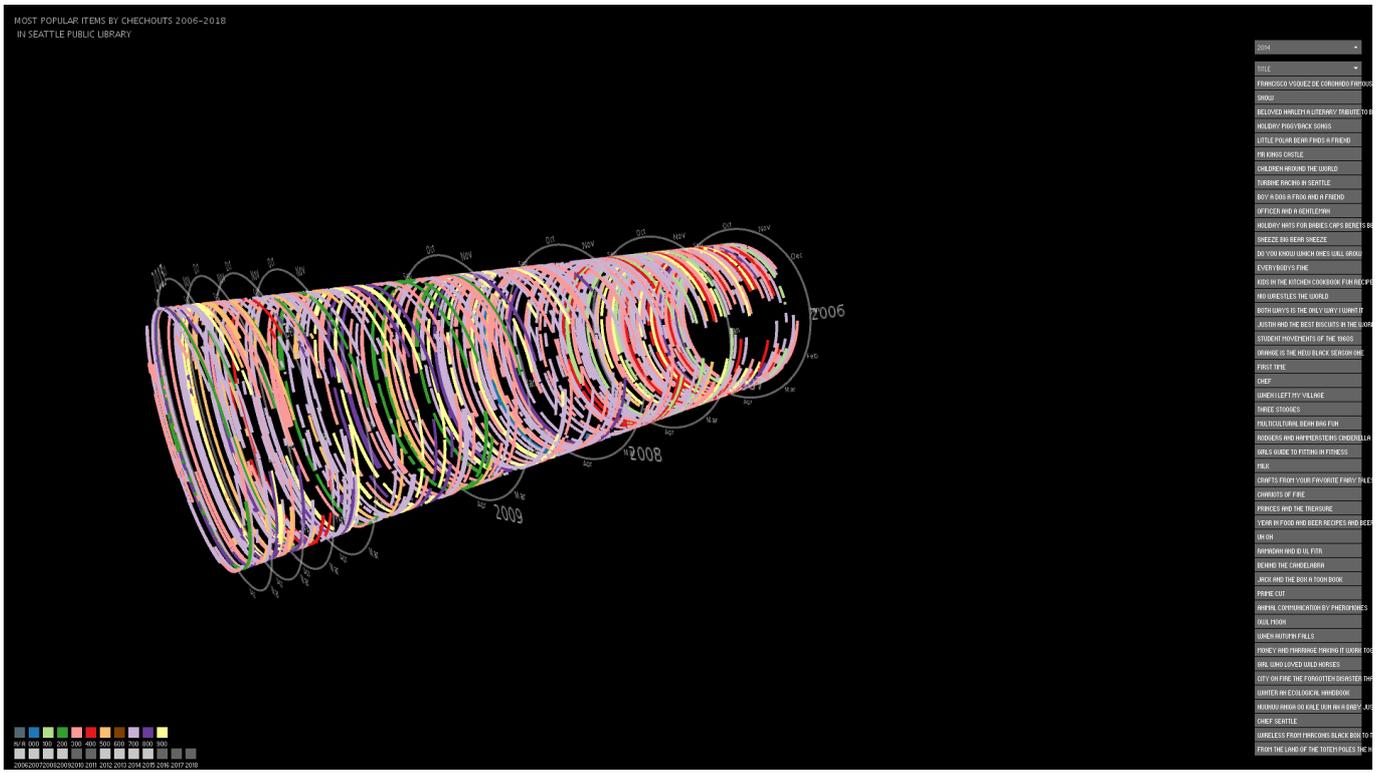
I used different colors for the Dewey classes. However, I wasn't using x and y-axis for any information. For each record, I just generated a random point inside the circle.

I decided to draw the checkout records as arcs around the circle. The circle starts from 0 to 360 degrees, and the degree can be mapped to month, from January to December. I also used brighter colors for different Dewey classes. Below are some screenshots.









Final Results & Analysis

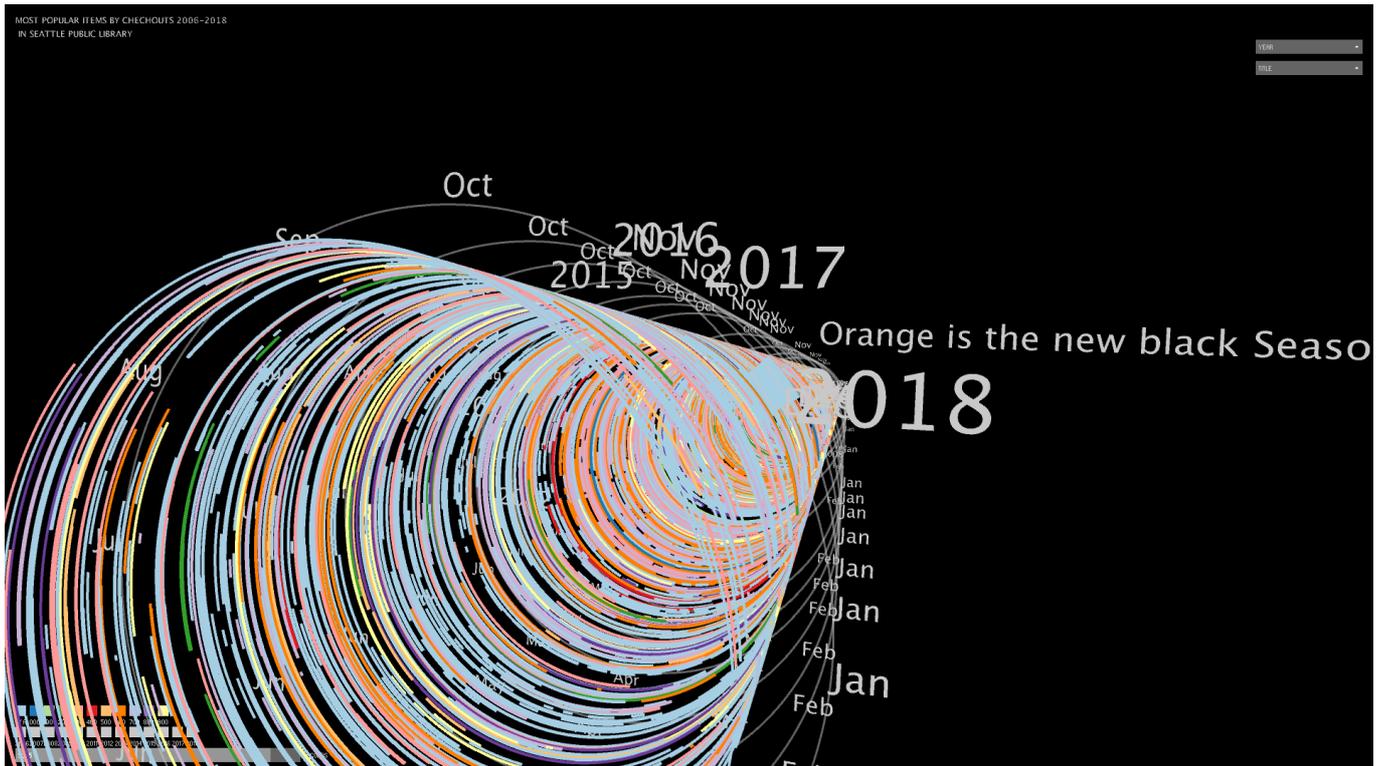
From the visualization, we can clearly see that the checkout times dropped year by year, which matches the conclusions that some of my peers draw. Also, a large portion of the popular checkouts is from non-dewey class, which means they don't have a Dewey class associated with them. I also think that Peasycam and ControlP5 are two very good libraries and it really saved me a lot of time when developing this application.

Future Improvements

I want to add hover text on each record for future improvements. I'm also considering adding a presentation mode which moves the camera slowly to show each year's titles.

Update * 2/19/2019

I made some updates to add the hover text using a slider, since I found it rather hard to check if the mouse is on an arc. Here are some screenshots.



The full list of updates:

- make text cleared using `textMode()`
- add a slider to let the user see titles of the checkout