AN ANALYSIS OF CONSUMER SPENDING ON ENTERTAINMENT

Presented By

Jocelyn Allen
Kathleen Fox
Kyle Lee
Mohammad Ronosentono

Methodology

- Used API to retrieve data and create analysis from the Bureau of Labor Statistics (BLS) 2017 Consumer Expenditure Survey

- About the Consumer Expenditure Survey
  - Nationwide Household Survey conducted by BLS
  - Conducted to determine how Americans spend their money
  - Consists of estimates from two separate surveys: An Interview and Diary Survey
  - Recall survey of a period of 3 months or longer spent on various frequently purchased items.
  - Entertainment broken into five categories: Pets, Toys, TV and Audio, Fees and Admissions, and Other

Research Questions

- What are the trends from 2013-2017 for the overall entertainment category in total consumer expenditure?
  - By age
  - By race
  - By region

- Which entertainment subcategories have the highest consumer expenditure?
  - By age
  - By race
  - By region
-
- What is the % of average total expenditure spent on entertainment?
  - By age
  - By race
  - By region

Data Exploration

Consumer Expenditure Survey

Survey Overview The following is a sample format description of the Consumer Expenditure Survey series identifier:

```
                          1         2
                12345678901234567890
Series ID       CXUMENBOYSLB0101M
Positions       Value               Field Name
1-2             CX                  Prefix
3               U                   Seasonal Adjustment Code
4-11            MENBOYS             Item Code
12-15           LB01                Demographics Code
16-17           01                  Characteristics Code
18              M                   Process Code
```

To assist you in formatting series IDs, access any of the following for a list of codes and their corresponding titles:

- Item Codes
- Demographics Codes
- Characteristics Codes
- Process Codes

Data Exploration (Cont.)

```
import requests
import json
import pprint as pp
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
headers = {'Content-type': 'application/json'}
data = json.dumps({"seriesid": ['CXUENTRTAINLB0901M',
                                'CXUENTRTAINLB0903M',
                                'CXUENTRTAINLB0904M',
                                'CXUENTRTAINLB0905M',
                                'CXUENTEROTHLB0901M',
                                'CXUENTEROTHLB0903M',
                                'CXUENTEROTHLB0904M',
                                'CXUENTEROTHLB0905M',
                                'CXUFEESADMLB0901M',
                                'CXUFEESADMLB0903M',
                                'CXUFEESADMLB0904M',
                                'CXUFEESADMLB0905M',
                                'CXUPETSLB0901M',
                                'CXUPETSLB0903M',
                                'CXUPETSLB0904M',
                                'CXUPETSLB0905M',
                                'CXUTOYSLB0901M',
                                'CXUTOYSLB0903M',
                                'CXUTOYSLB0904M',
                                'CXUTOYSLB0905M',
                                'CXUTVAUDIOLB0901M',
                                'CXUTVAUDIOLB0903M',
                                'CXUTVAUDIOLB0904M',
                                'CXUTVAUDIOLB0905M',
                                'CXUTOTALEXPLB0901M',
                                'CXUTOTALEXPLB0903M',
                                'CXUTOTALEXPLB0904M',
                                'CXUTOTALEXPLB0905M'
                               ],
                    "startyear":"2017",
                    "endyear":"2017",
                    "catalog":True,
                    "calculations":True,
                    "annualaverage":True,
                "registrationkey":"a3f26c081d1546a58b1461108c6f62f9"})
p = requests.post('https://api.bls.gov/publicAPI/v2/timeseries/data/', data=data, headers=headers)

json_data = json.loads(p.text)

pp.PrettyPrinter(indent=2)
output = open("CXENTERTAIN" + '.json','w')
pp.pprint(json_data, stream=output)
output.close()
```

```
list1 = []
tlist1 = []
tlist2 = []


if json_data['status'] == "REQUEST_SUCCEEDED":
    for series in json_data['Results']['series']:
        seriesId = series['seriesID']
#         print(series['catalog']['item'])
        series_title = series['catalog']['series_title']
        titleList = series_title.split('by Race:', 1)
        tlist1.append(titleList[0])
        tlist2.append(titleList[1])
        #print(series_title)
        for item in series['data']:

            year = item['year']
            period = item['period']
            value = int(item['value'])
            list1.append(value)
            dict1 = {''}
        #print("Year:" + year + " Value:" + value)

else:
    print(json_data['status'])
```

```
df = pd.DataFrame(np.column_stack([list1, tlist1, tlist2]),
                                  columns=['value', 'Category', 'Race'])
df.value = pd.to_numeric(df.value, errors='coerce')
```

Data Exploration (Cont.)

```python
import requests
import json
import pandas as pd
import api_keys
import time

class BLSservice:
    headers = {'Content-type': 'application/json'}

    #For Current Year data
    def __init__(self, series_id):
        self.series_id = series_id
        curr_time = time.localtime()
        self.start_year = str(curr_time.tm_year)
        self.end_year = str(curr_time.tm_year)
        self.catalog = True
        self.calculations = True
        self.annualaverage = True

    #For Multiple Year data with options
    def __init__(self, series_id, start_year, end_year, catalog= True, cal
        self.series_id = series_id
        self.start_year = start_year
        self.end_year = end_year
        self.catalog = catalog
        self.calculations = calculations
        self.annualaverage = annualaverage

    def get_data(self):
        try:
            self.data = json.dumps({"seriesid": self.series_id,
                        "startyear":self.start_year,
                        "endyear":self.end_year,
                        "catalog":self.catalog,
```

In [1]:
```python
import requests
import json
import pprint as pp
import pandas as pd
from laborstat import BLSservice

import numpy as np
import matplotlib.pyplot as plt
```

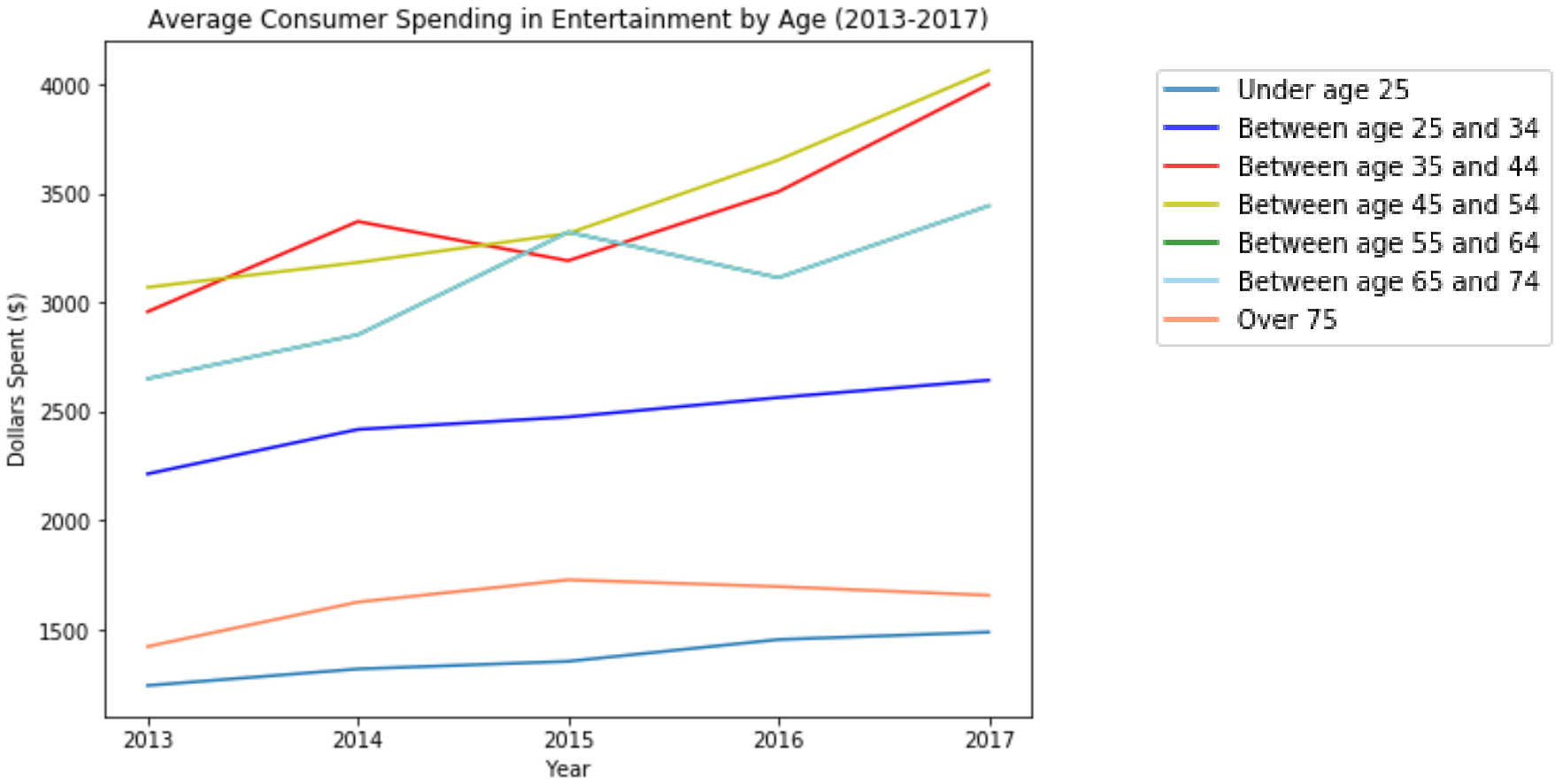
Entertainment Total Spending based on Age

In [2]:
```python
overall_on_age = BLSservice([ 'CXUENTRTAINLB0401M',
                              'CXUENTRTAINLB0402M',
                              'CXUENTRTAINLB0403M',
                              'CXUENTRTAINLB0404M',
                              'CXUENTRTAINLB0405M',
                              'CXUENTRTAINLB0406M',
                              'CXUENTRTAINLB0407M',
                              'CXUENTRTAINLB0408M',
                              'CXUENTRTAINLB0409M'
                            ], '2013', '2017')
```

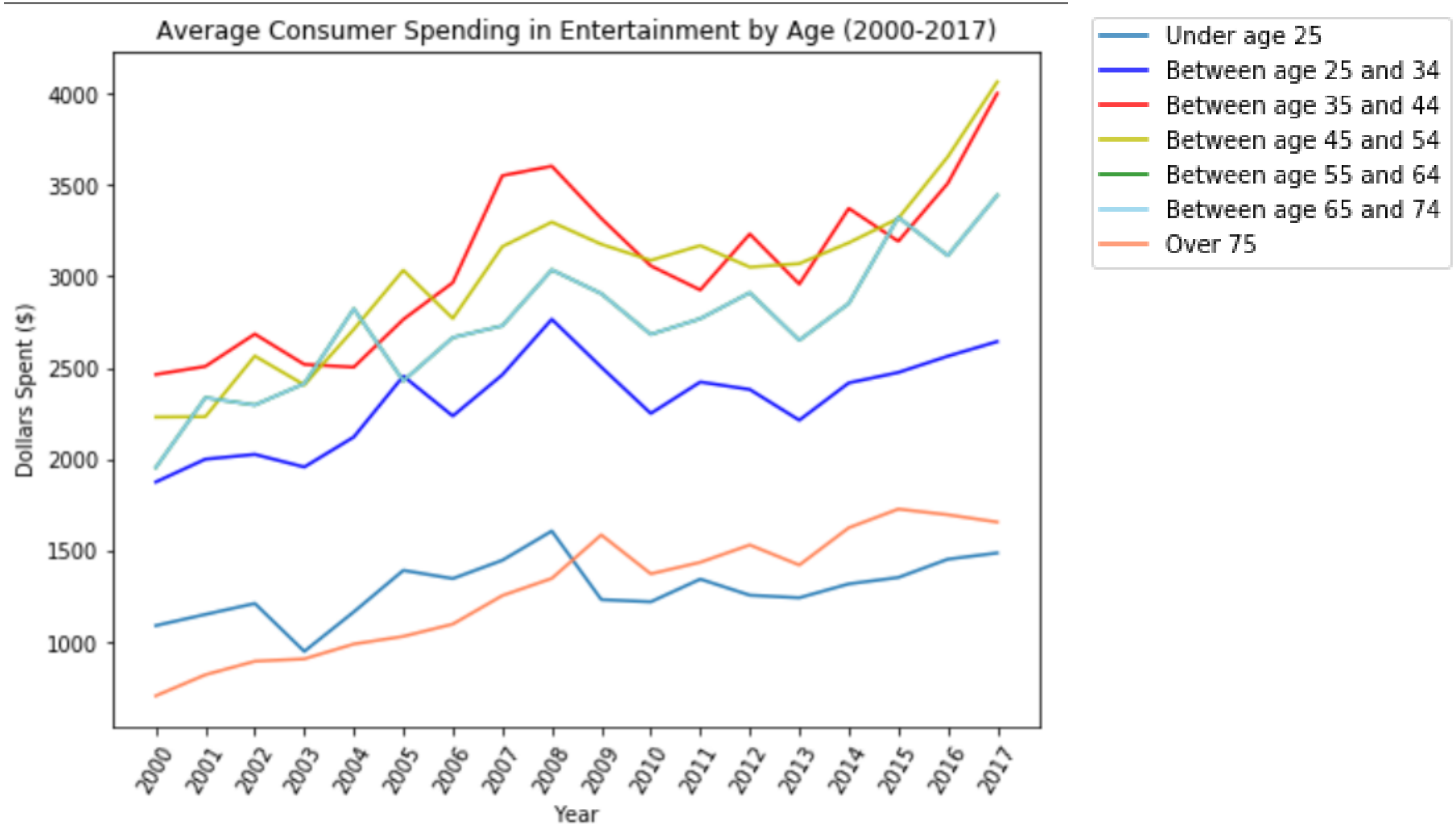
In [3]:
```python
overall_on_age.get_data()
overall_on_age.create_dataframe()
```

RESULTS: TRENDS BY AGE

What are the trends from 2013-2017 for the overall entertainment category in total consumer expenditure?

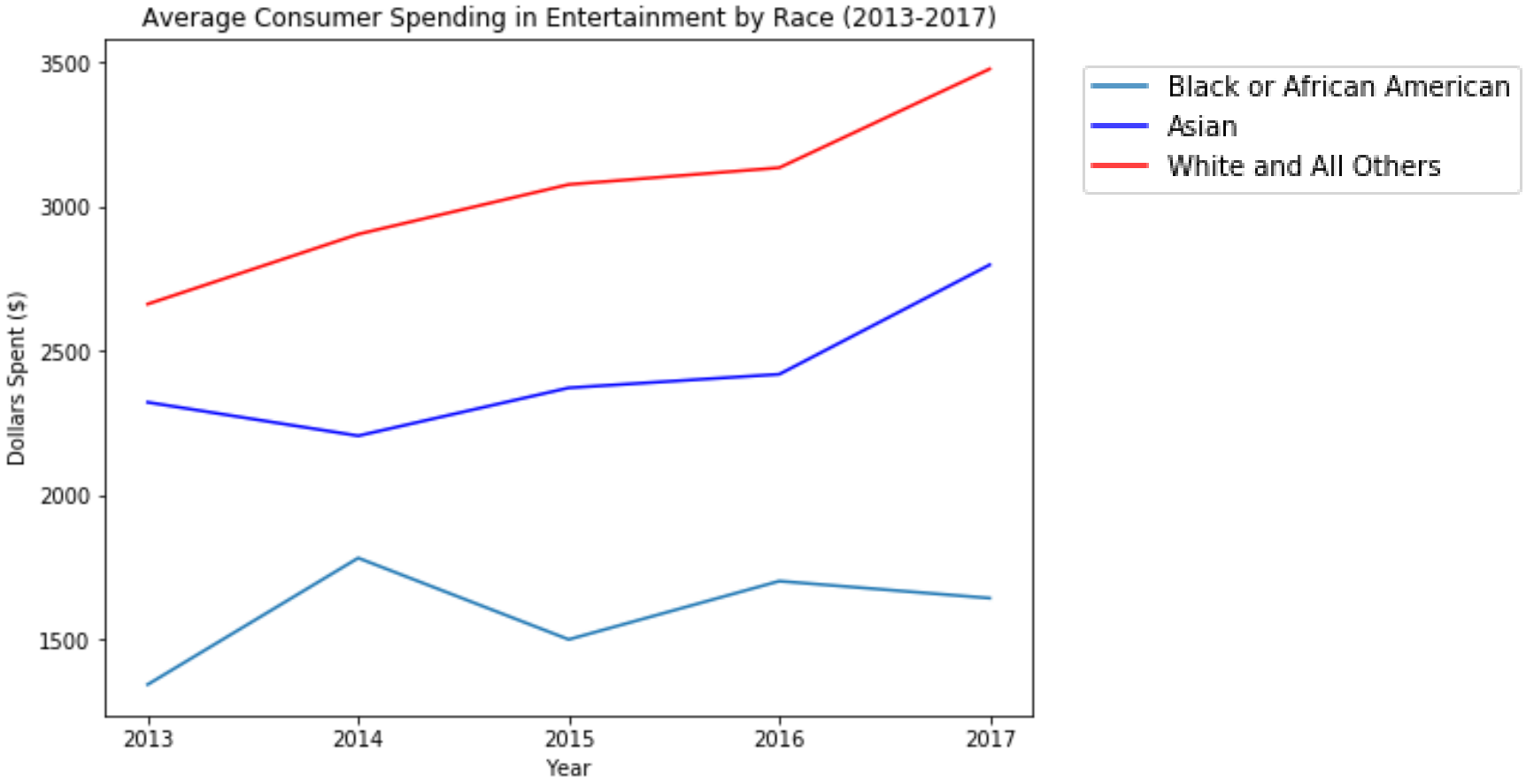


RESULTS: TRENDS BY AGE (Comparison)

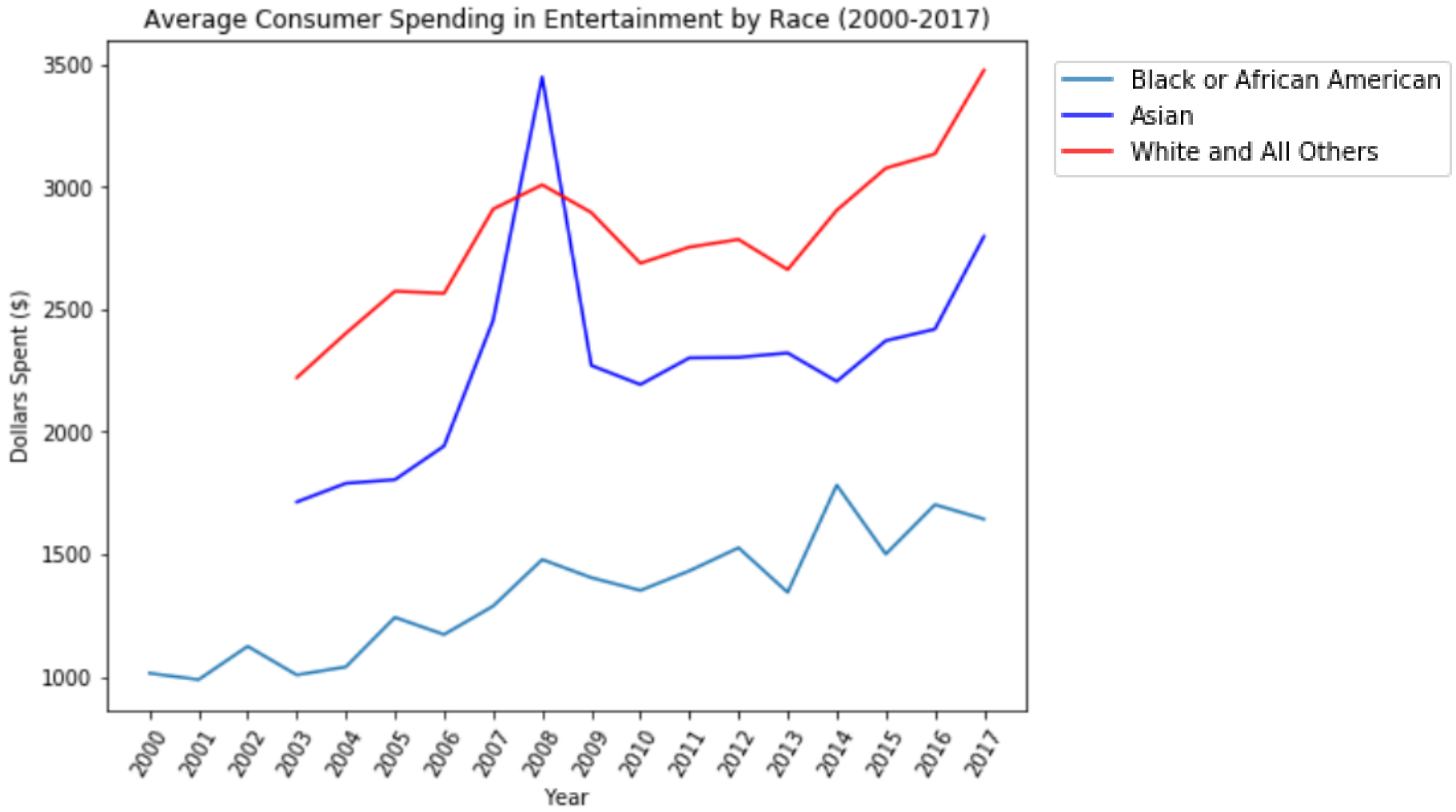


RESULTS: TRENDS BY RACE

What are the trends from 2013-2017 for the overall entertainment category in total consumer expenditure?

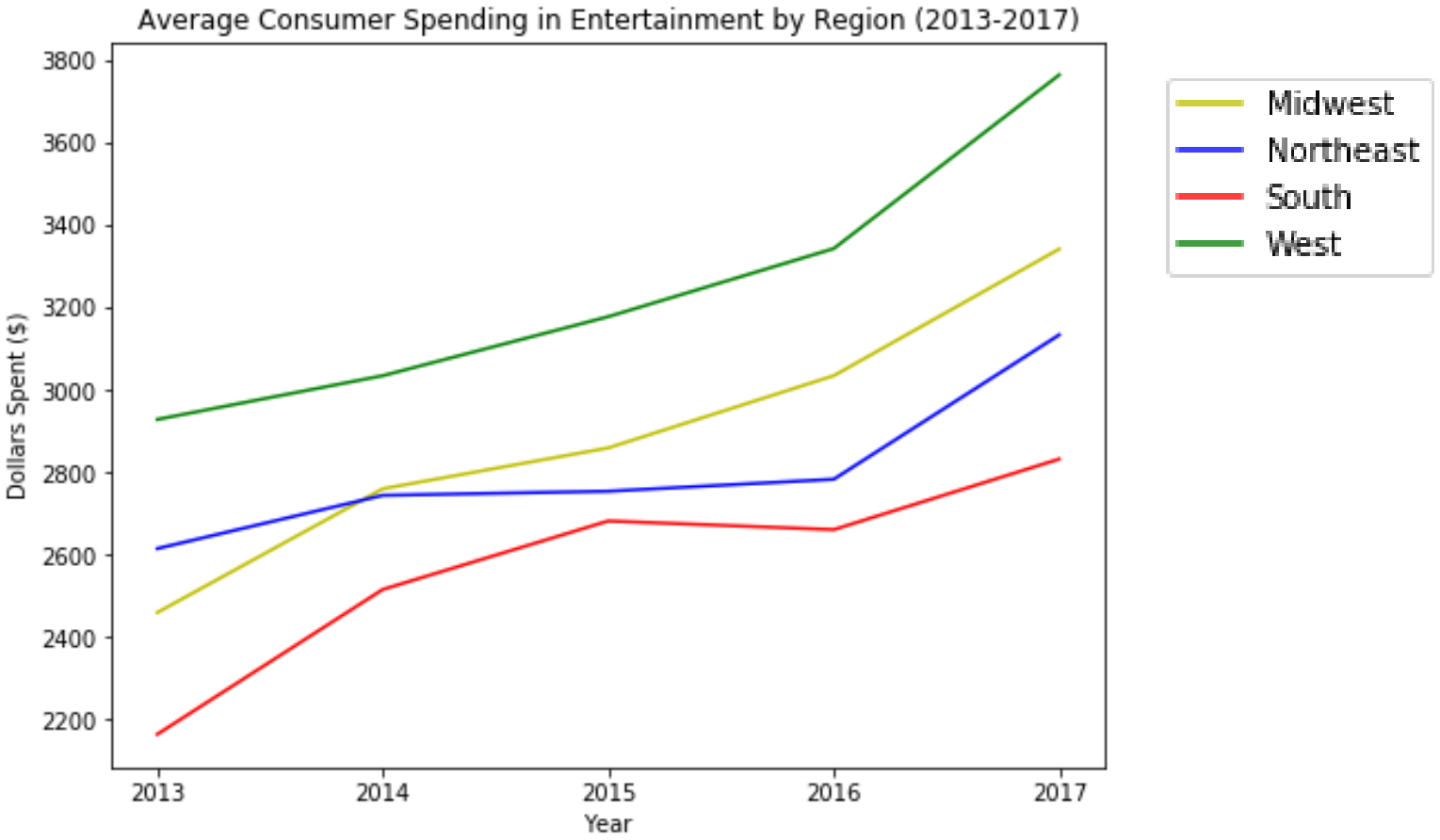


RESULTS: TRENDS BY RACE (Comparison)

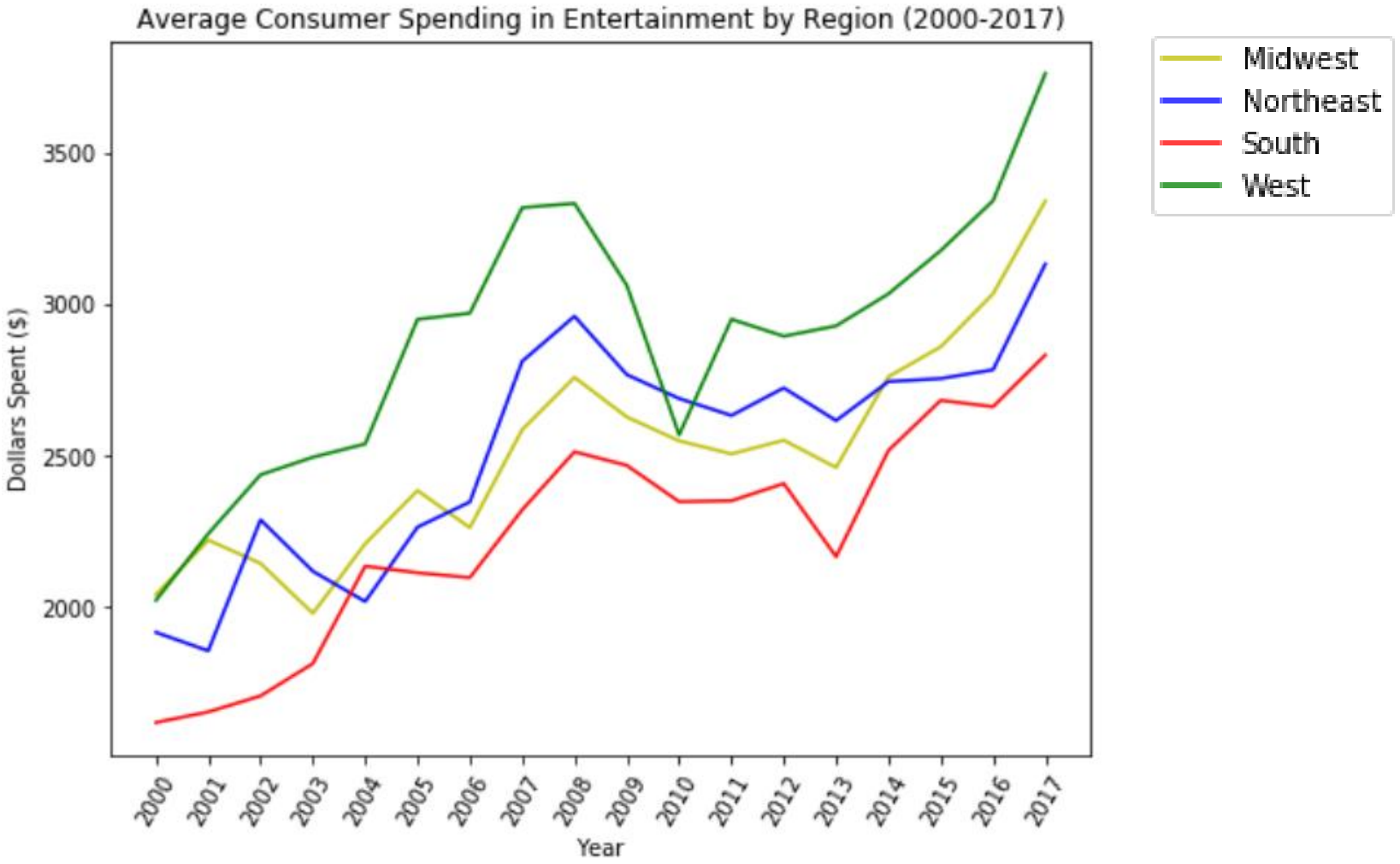


RESULTS: TRENDS BY REGION

What are the trends from 2013-2017 for the overall entertainment category in total consumer expenditure?

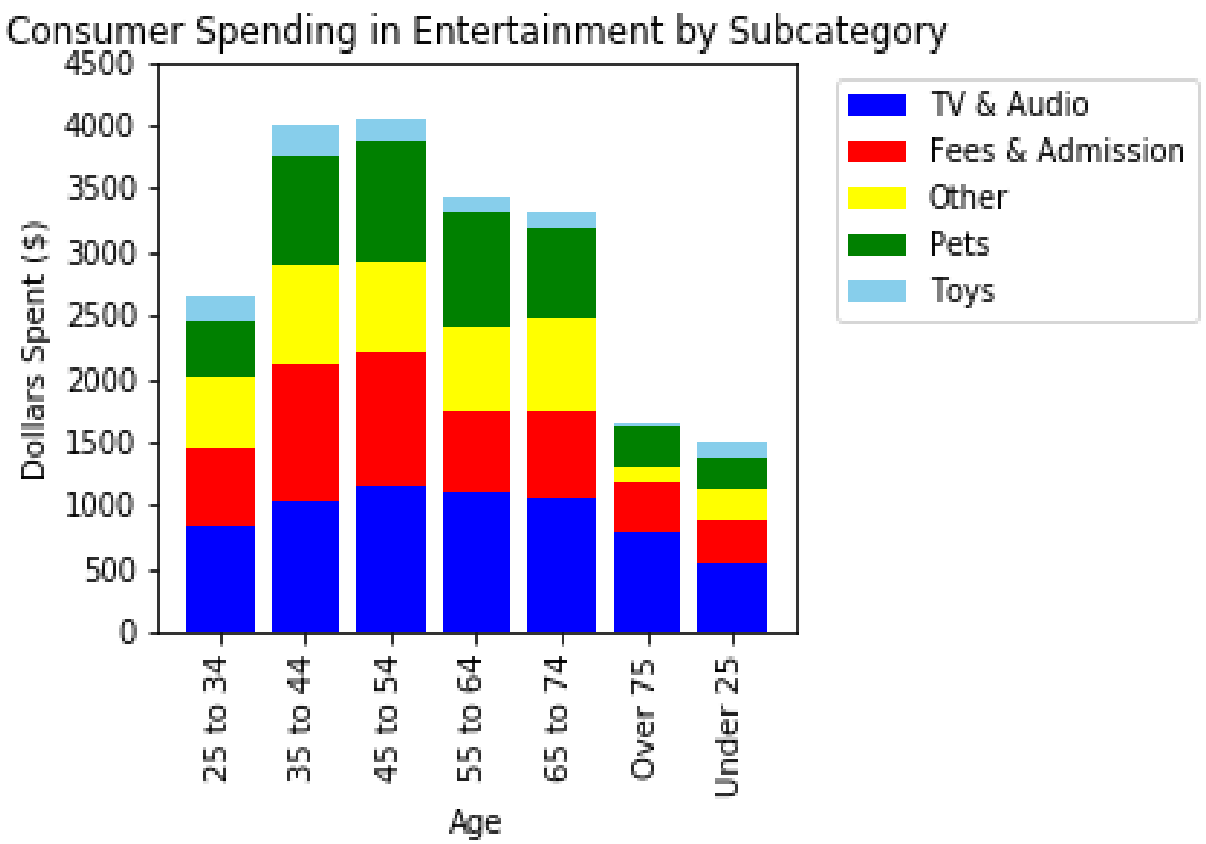


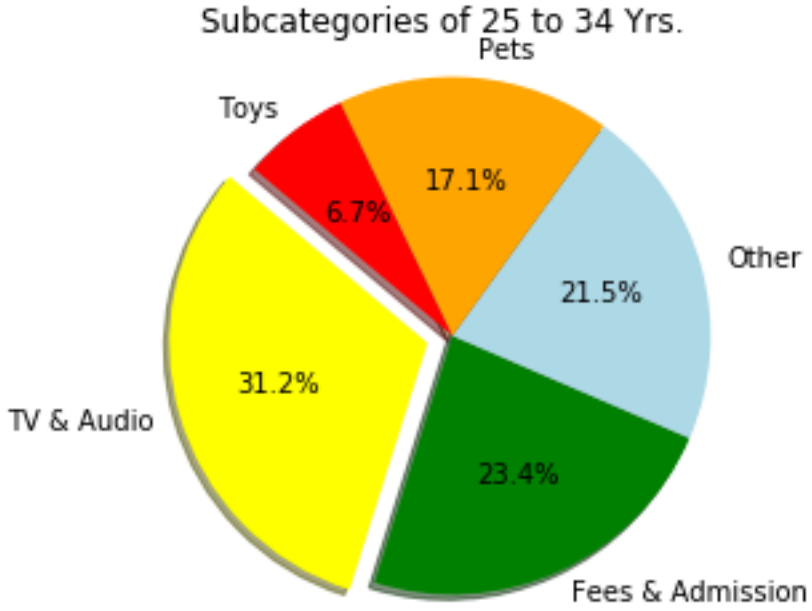
RESULTS: TRENDS BY REGION (Comparison)



RESULTS: AGE

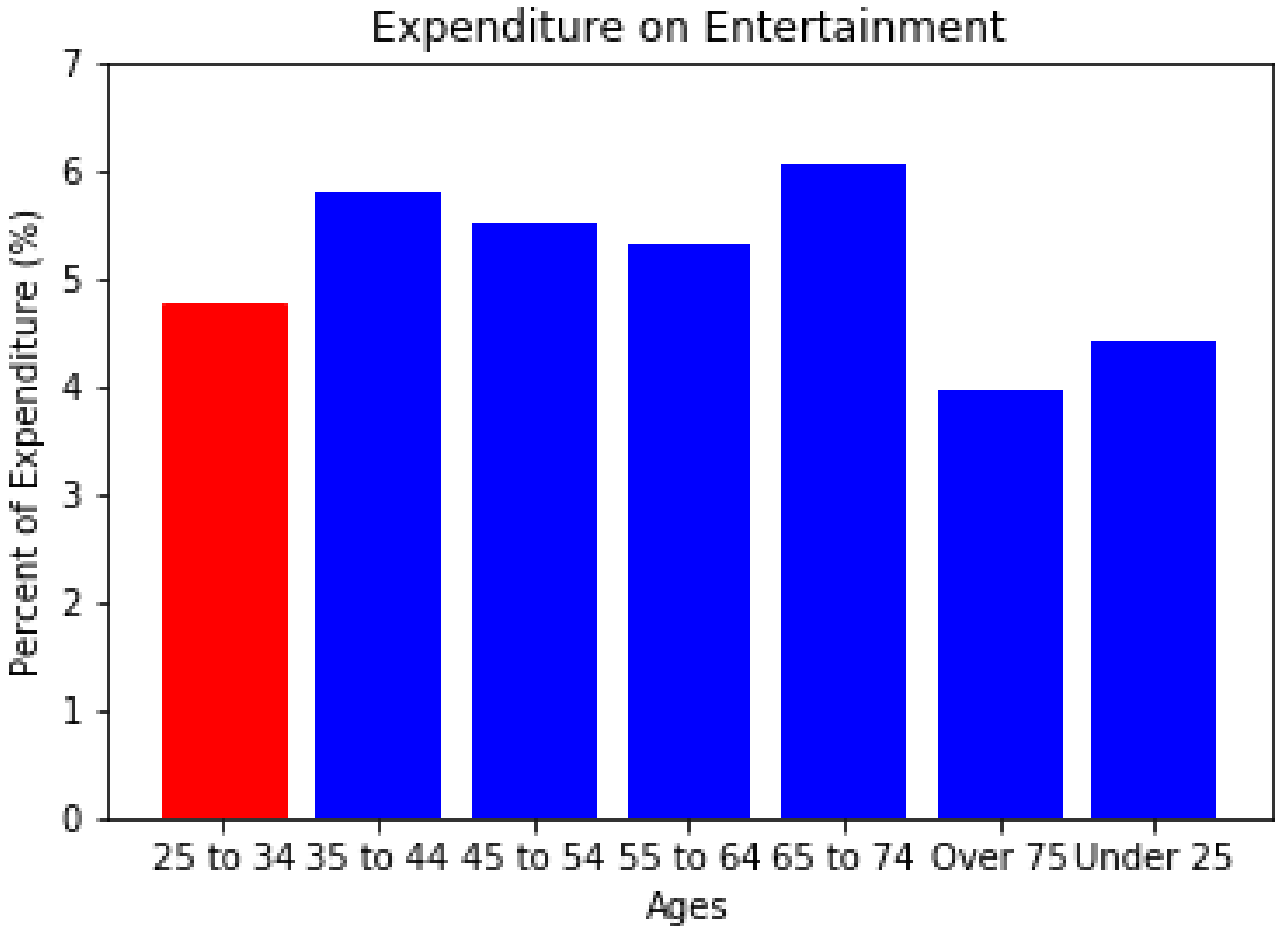
Which entertainment subcategories have the highest consumer expenditure by age?





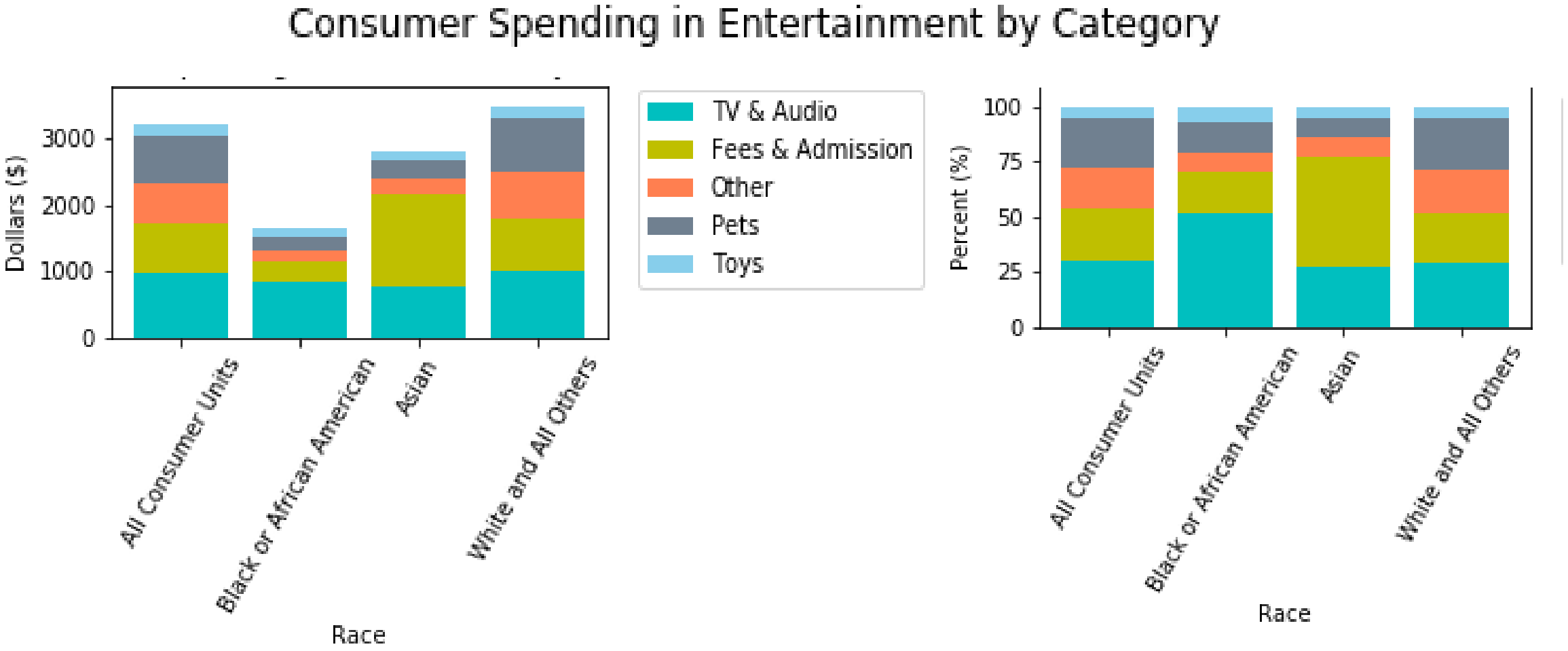
RESULTS: AGE

What is the % of average total expenditure spent on entertainment by age group?



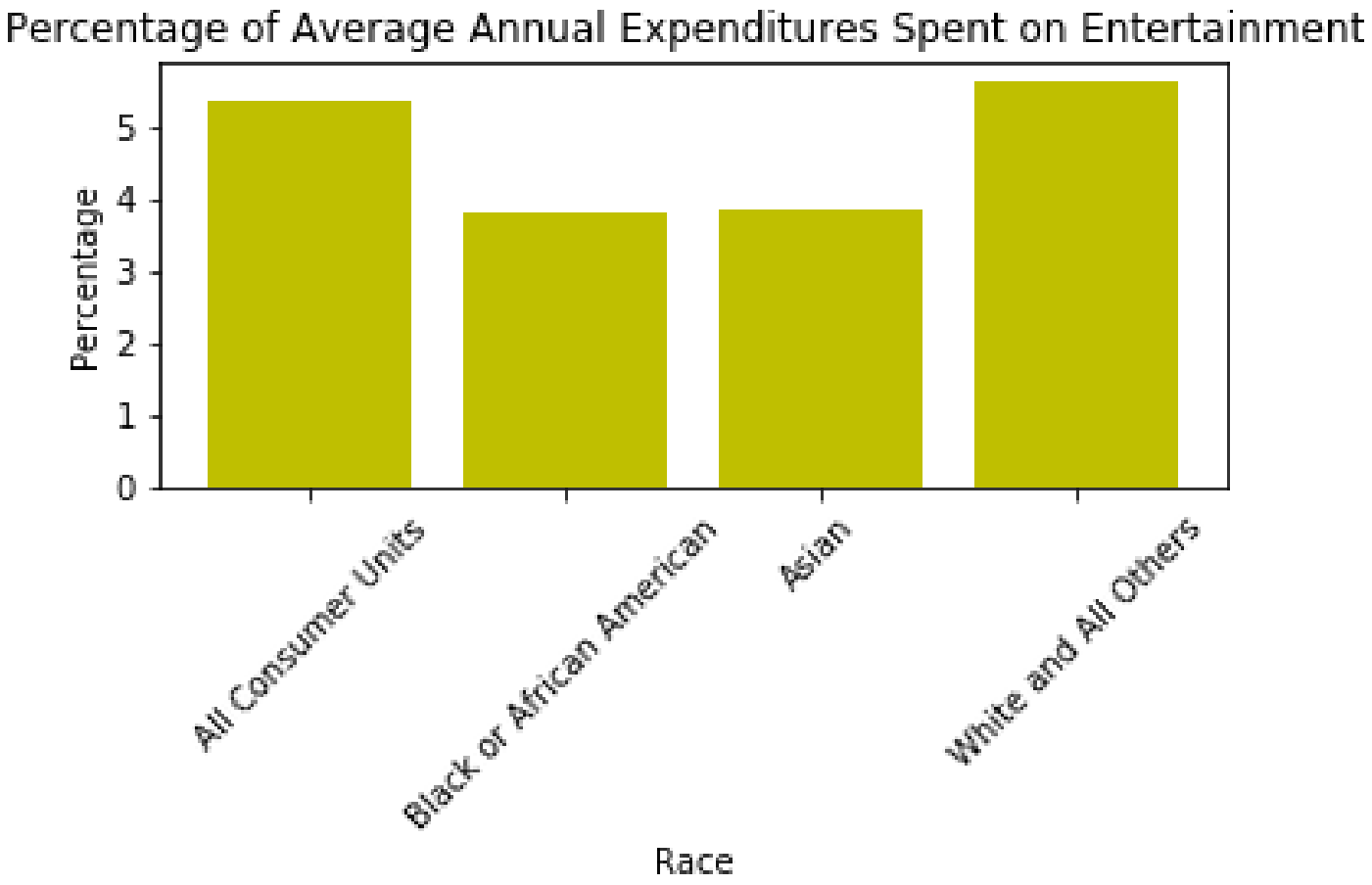
RESULTS: RACE

Which entertainment subcategories have the highest consumer expenditure by race?



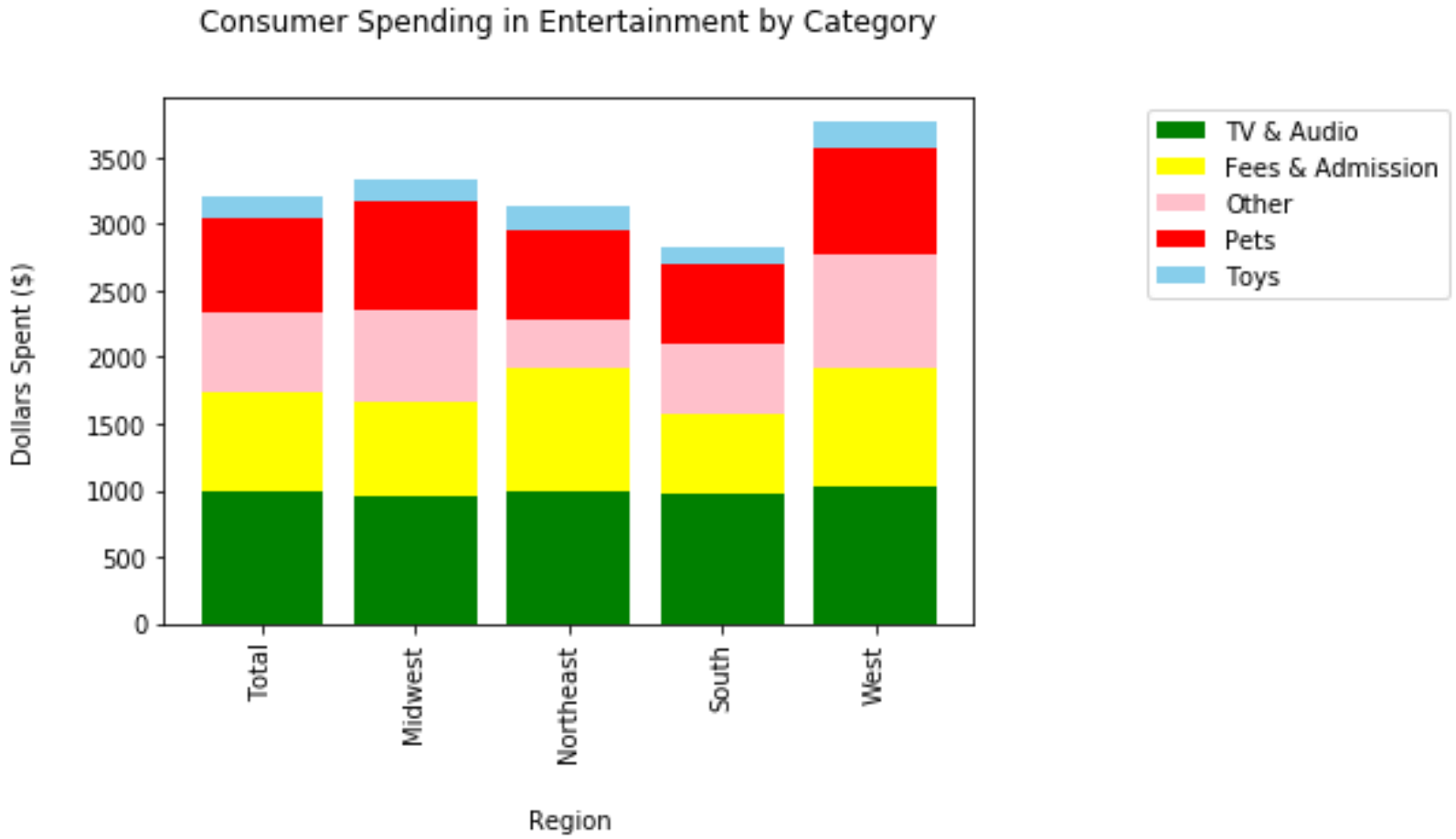
RESULTS: RACE

What is the % of average total expenditure spent on entertainment by race?



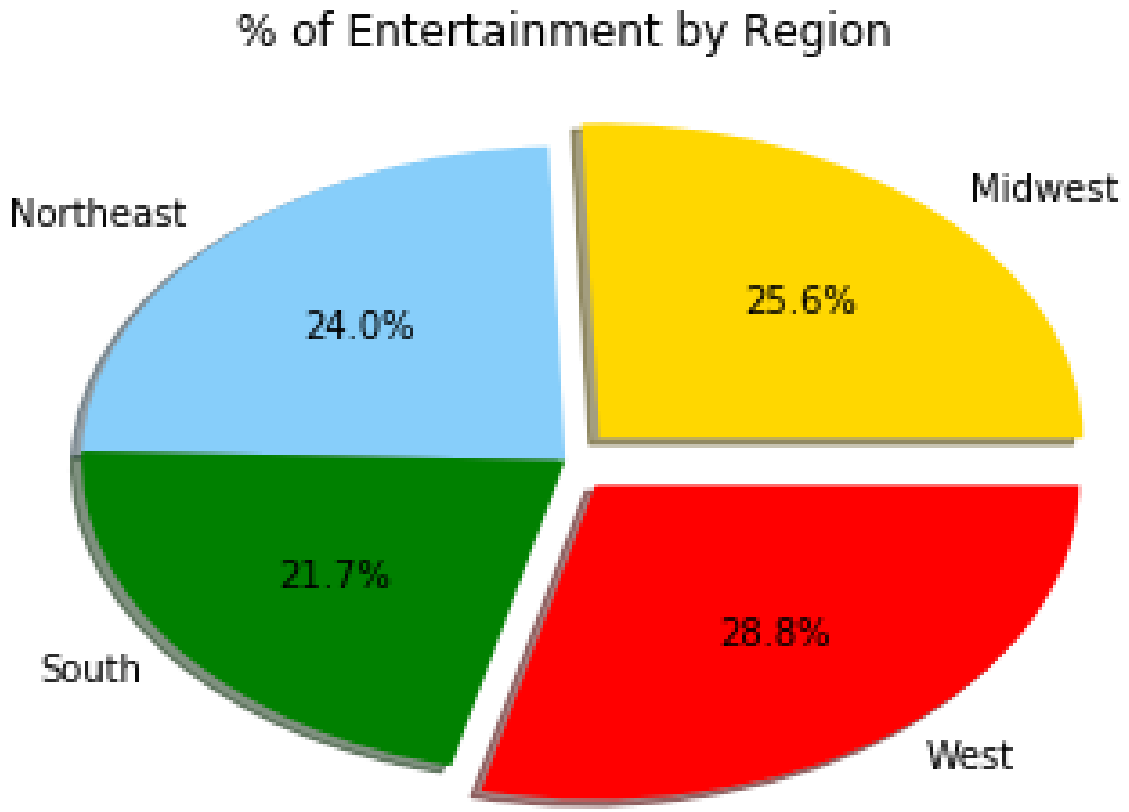
RESULTS: REGION

Which entertainment subcategories have the highest consumer expenditures by region?



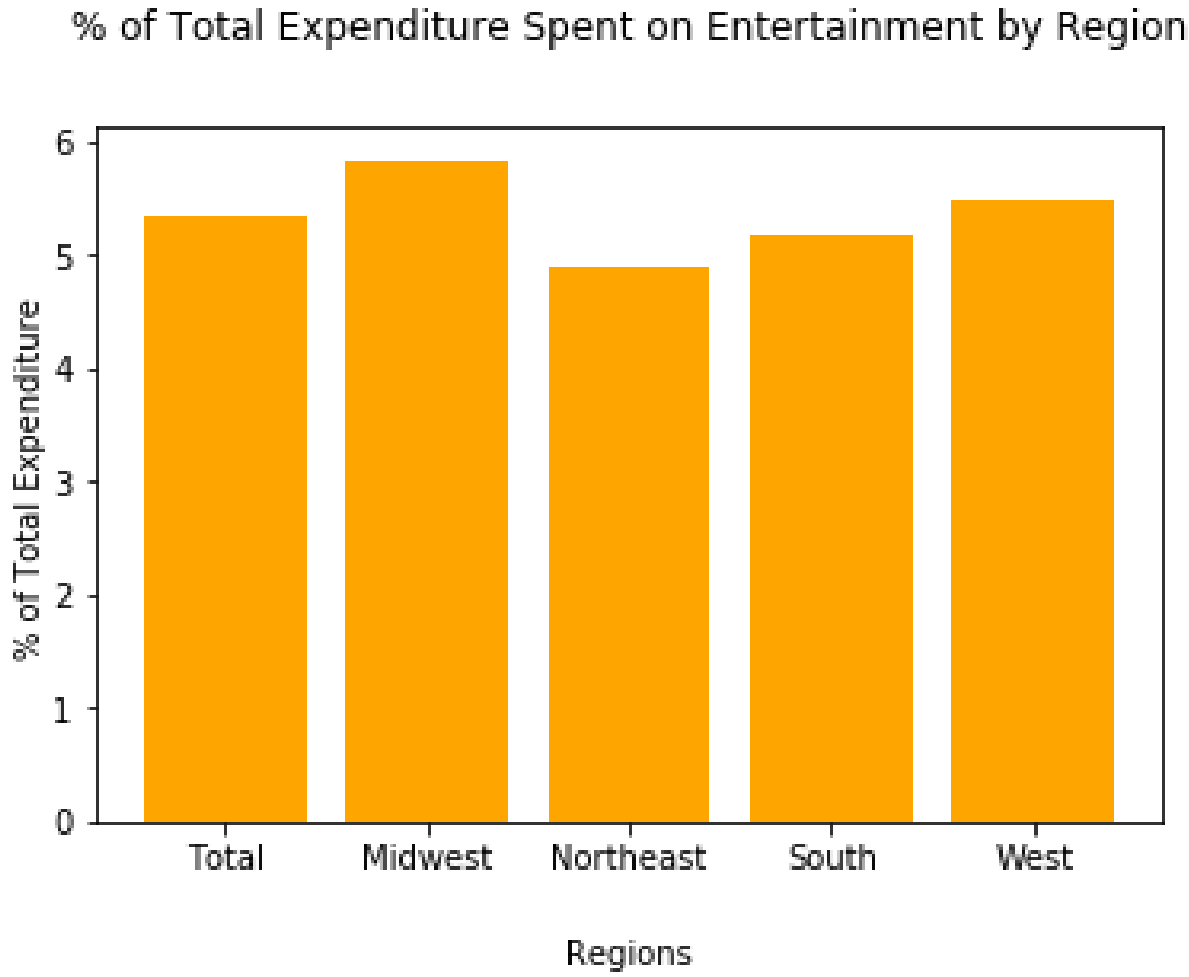
RESULTS: REGION

Which entertainment region has the highest consumer expenditure for entertainment?



RESULTS: REGION

What is the % of average total expenditure spent on entertainment by region?



CONCLUSIONS

- 35 to 44 and 45 to 54 year olds both saw sharp increases in entertainment spending from 2016 to 2017
- Black/African American and Asians saw increases in entertainment spending from 2016-2017 while White remained relatively flat.
- All four geographical regions saw increases in entertainment spending from 2016-2017.
- 35 to 44 and 45 to 54 year olds have near identical spending habits.
- 65 to 74 year olds spend the most percent of their total expenditure on entertainment.
- Asians spent the largest percent of their entertainment expenditures on fees and admissions.
- White people spent the highest percentage of total expenditures on entertainment.
- The West Region spent the most money on Entertainment.
- The Midwest Region spent the largest amount of their total expenditures on on Entertainment.

AREAS FOR FURTHER STUDY

- How does the US entertainment industry compares to the worldwide entertainment industry when looking at demographics?
- Looking at consumer spending overall, not limit it to the entertainment category
- Factors contributing to spending per category