Geena Davis Institute on Gender in Media

If she can see it, she can be it.™

Final Project Report Draft

25/04/2022

—

Sponsors

Meredith Conroy

Romeo Perez

Project Manager

Grace Yoon

Team Members

Dayong Wu (Team Leader)

Emmanouil Kritharakis

Yan Tong

Junfei Huang

Yuanli Wang

Abstract

Background

During this semester, our team worked on the **See Jane | Research Project Data Normalization** project. Our clients are Meredith Conroy and Romeo Perez,, who are members of the See Jane organization. This project aims to help the See Jane organization normalize, merge, and analyze their datasets related to the entertainment industry.

Both technical and non-technical prerequisites are needed to handle this project.

- Prior knowledge to the entertainment industry
- Sense of social responsibility of gender balance, race equality, diversity and inclusion, etc
- Comprehensive understanding of the codebooks
- Python Pandas, Numpy, Matplotlib (etc), and Excel skills are prerequisites

Motivation

- Create gender balance, foster inclusion and reduce negative stereotyping in family entertainment media.

Goal

- Normalize the datasets & refine the codebooks
- Analyze industry differences in representation in advertisements
- Summarize the change in representation in advertisements
- Detect the trends of representations in advertisements

Exploration

Before performing analysis on the datasets, our team has done a lot of data exploration work, which equips us with all necessary understandings about the datasets.

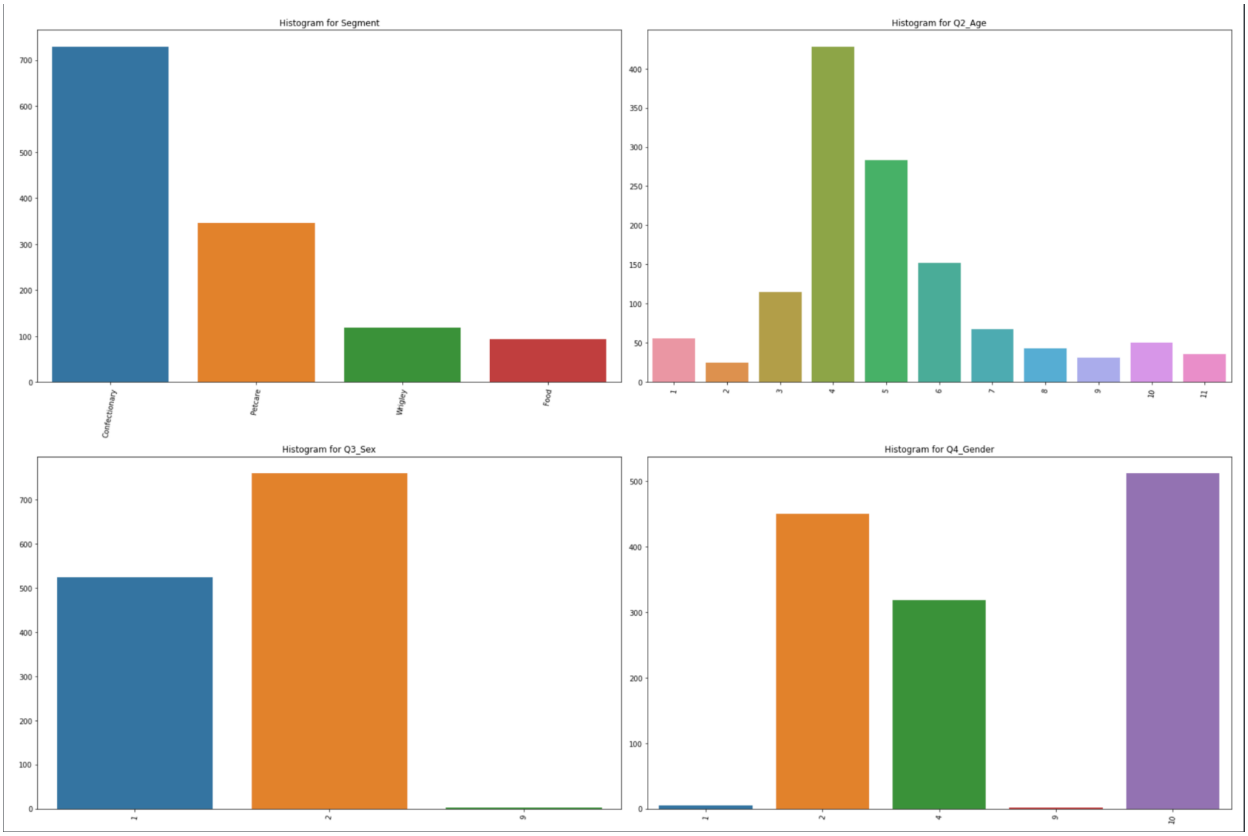
Through data exploration, we have developed a holistic comprehension of the meaning of each column and its values, different categories of the columns, the size of the dataset, etc. Below are our findings of some overarching columns:

- Q2: Age. Enter your best estimate of the character's age. If the character has multiple ages in the commercial, select "other" and provide details.
- Q3: Sex. Enter your best assessment of the character's "sex."
- Q4: Gender. Enter your best assessment of the character's gender performance, regardless of their sex. Masculinity refers to a set of stereotypical male traits and behavior, including assertiveness, being in control, aggression, an emphasis on physical strength, and sexual promiscuity. Femininity refers to a set of stereotypical female traits and behaviors, including passivity, an emphasis on being pleasing, gentleness, dependence, and an emphasis on caring and empathy. Hyper-masculinity and hyper-femininity are exaggerations of these gender performances (think of the typical roles played by Arnold Schwartzenegger and Marilyn Monroe, respectively). Gender queer and gender non-conforming describes for characters who do not fit conventional gender distinctions, rather, they identify with neither or a combination of masculinity and femininity. Code the character as "feminine" or "masculine" unless cued otherwise.
- Q5: LGBTQ. A character's sexuality is determined by his/her apparent enduring attraction (emotional, sexual, romantic) to men, women, or both sexes. Code the character as "heterosexual" unless cued otherwise.

Analysis

Result for Question 1: Are there industry differences in representation in advertisements?

First we have an observation for entire dataset
1. For Segment column, we can find the 'confectionary' accounts for half of the total data volume of Segment column
2. For Q2_Age column, most of values are Q2_Age = 4, which is the age of 20-29 year olds
3. For Q3_Sex column, number of male is more than female, and there is a little people don't tell their sex
4. For Q4_Gender column, we find that value 'Not Applicable' is the largest, and then is the 'Masculine' and 'Feminine'. Also, there are a few 'Hyper-Masciline' and 'Can't Tell'



Since “industry” refers to the “Segment” column, and “representation” refers to all the question columns. We find the industry differences among 'Q2_Age', 'Q3_Sex' and'Q4_Gender' columns, since they are common statistical variables

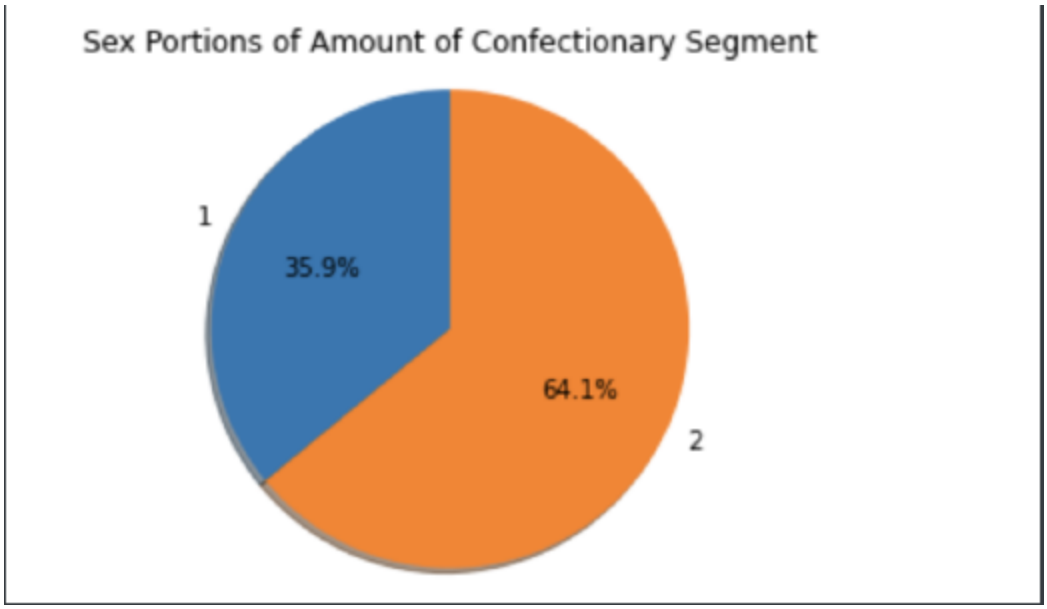
Observation of industry differences for Q2_Age:

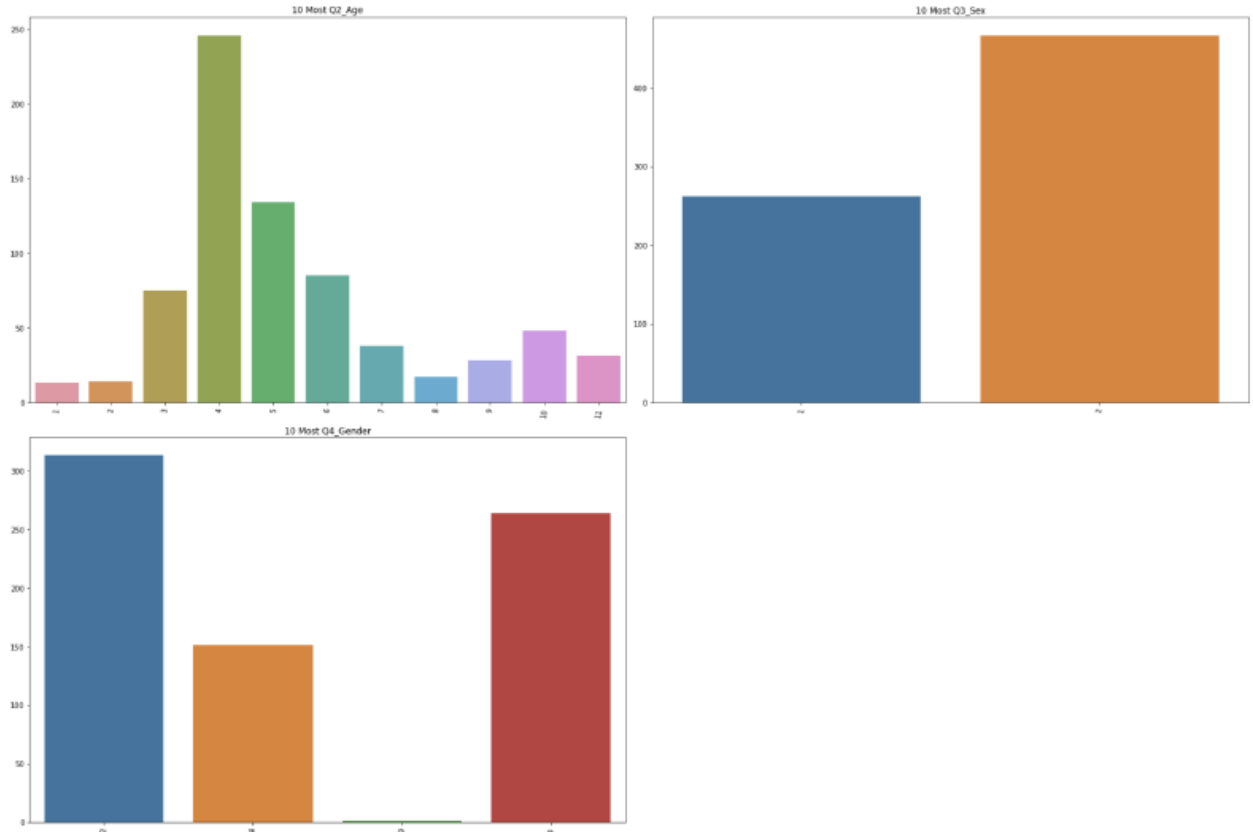
1. We find that among the Q2_Age=1, the value of Petcare industry is the highest
2. With the increasing of the age, more people are in the "Confectionary" industry, but when Q2_Age = 8 the Petcare is greater than Confectionary
3. By the statistics, we observe that the largest amount of data are in the "Confectionary" industry
4. And someone don't tell or are unwilling to disclose the information of age Observation of industry Differences for Q3_Sex:
1. We find that among the Q3_Sex = 1 which is "Female", the value of "Confectionary" is the highest, and the "Food" is the smallest
2. We find that among the Q3_Sex = 2 which is "Male", the value of "Confectionary" is the highest, and the "Food" is the smallest
3. We find that among the Q3_Sex = 9 which is "Can't tell", there is only Petcare
4. By the statistics, we observe that the largest amount of data are in the "Confectionary" industry
5. And some of people can not tell their sex

Observation of industry differences for Q4_Gender:

1. We find that among the Q4_Gender = 1 which is "Hyper-Masculine", is in either Petcare or Wrigley
2. We find that among the Q4_Gender = 2 which is "Masculine", the "Confectionary" is much higher than other industries, each industry has "Masculine"
3. We find that among the Q4_Gender = 4 which is "Feminine", the "Confectionary" is also higher than other industries, each industry has "Feminine"
4. By the statistics, we observe that the largest amount of data whatever their gender are in the "Confectionary" industry

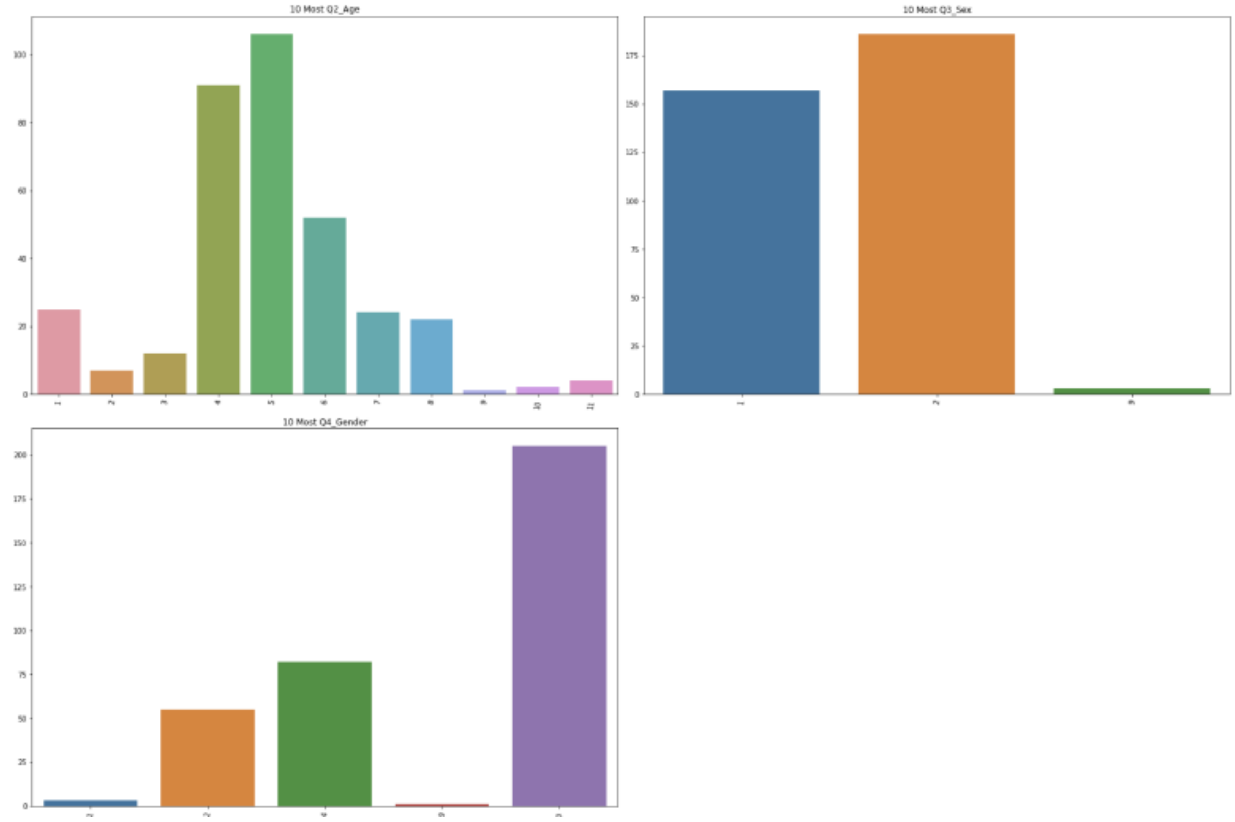
5. Large amount of people are "Not Applicable" of Q4_gender in "Confectionary", "Petcare" and "Food"
Then try different way to answer the question: separating four different segments to different dataframes and analyze.
Observation of industry differences for confectionary dataframe:
1. We find that for Confectionary, the 20-29 age group(4) has the largest proportion, the second one is group(5)
2. Through Q3_Sex and Q4_Gender, we find that the number of male is definitely greater than female and lots of people choose "Not Applicable" on Q4_gender
3. From barplot we find that for Confectionary segment, the amount of "Male(2)" is 64.1%, which is more than "Female(1)" (35.9%)
4. From four barplots for sex portions of different industry, we conclude that Confectionary has the largest difference in the number of male and female
5. We get the average value of Q2_Age for the Confectionary industry is 5.36(around 34 years old),the std value is 2.33
6. We can conclude that the the average age of the Confectionary industry is the highest



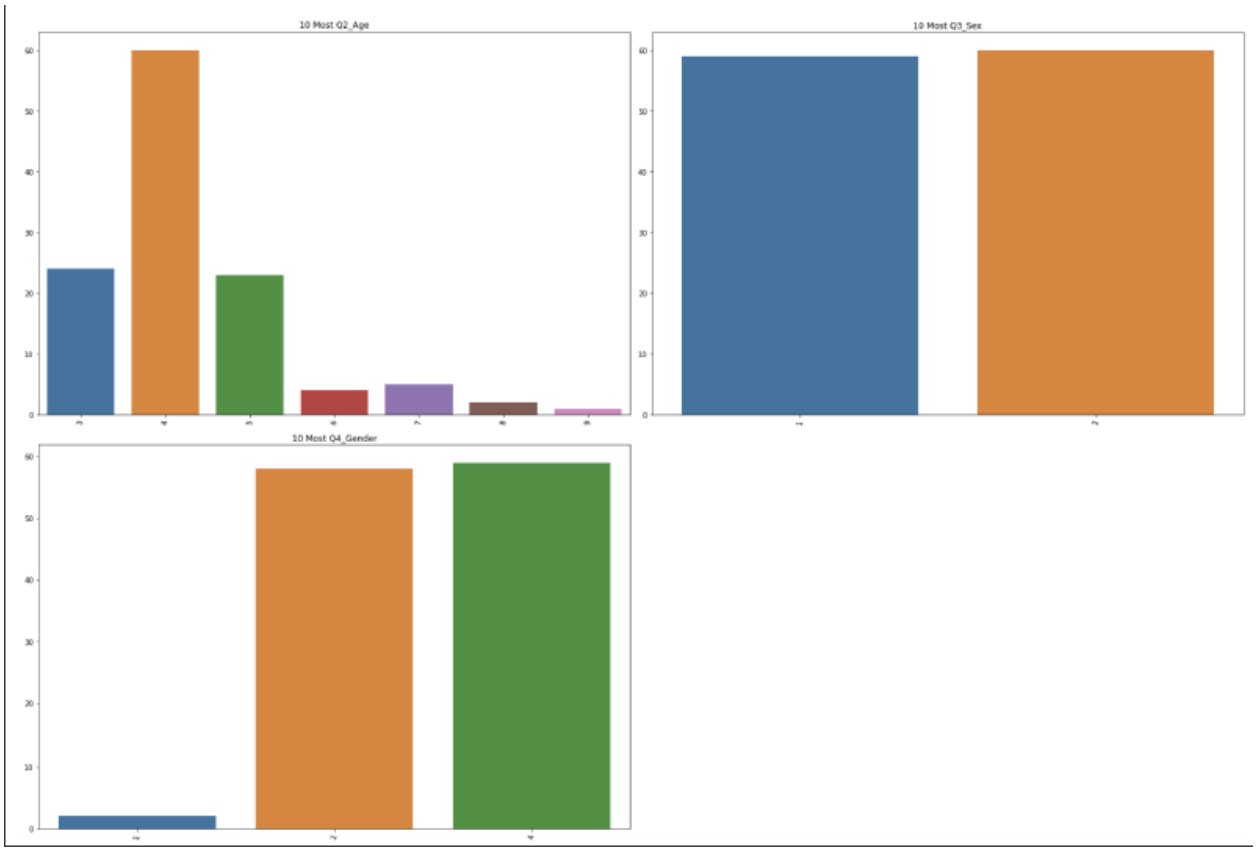


Observation of industry differences for Petcare dataframe:

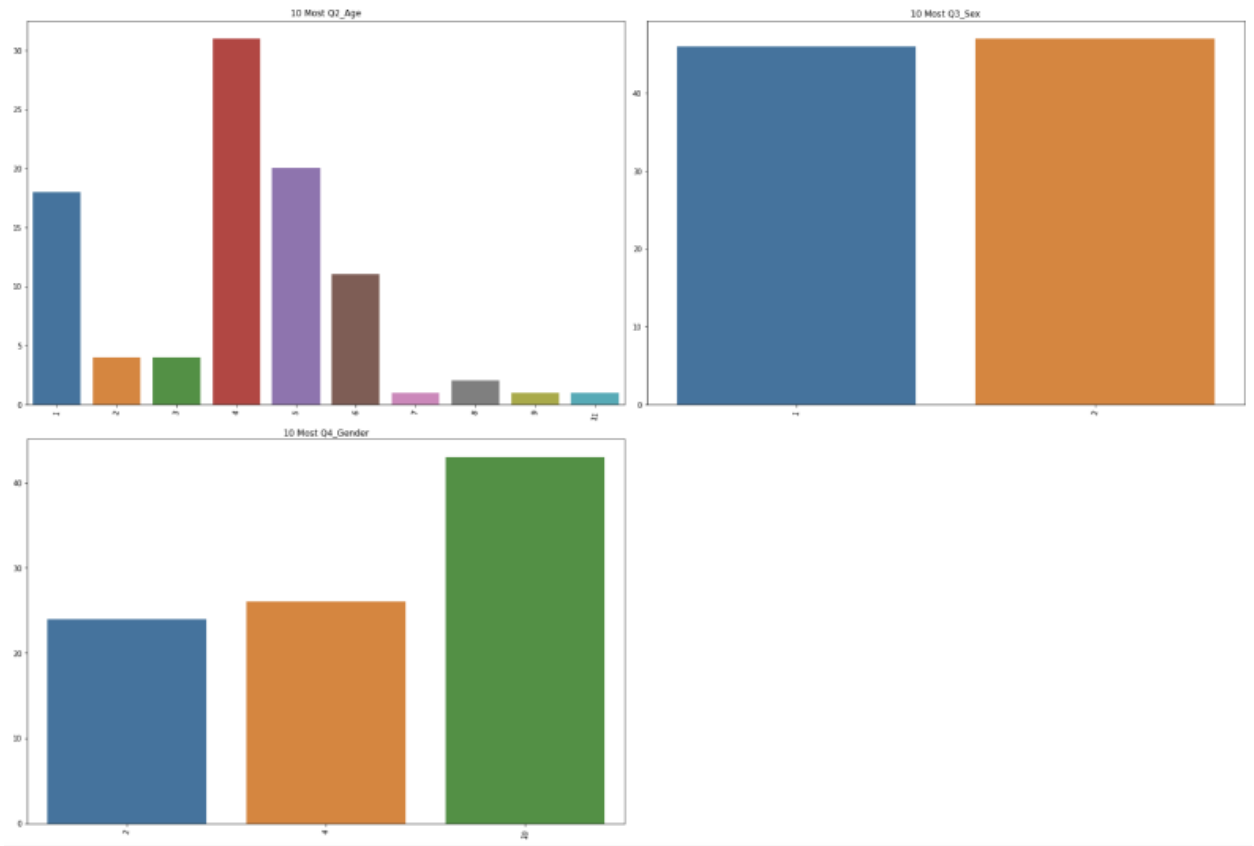
1. We find that for Petcare, the 30s age group(5) has the largest proportion, the second one is group(4)
2. From barplot we find that for Petcare, the amount of "Male(2)" is 53.8% which is the largest one, the amount of "Female(1)" is 45.4% and the amount of "Can't tell(9)" is 0.9%
3. We get the average value of Q2_Age for the Petcare industry is 4.91(around 30 years old), which is smaller than Confectionary industry, the std value is 1.83



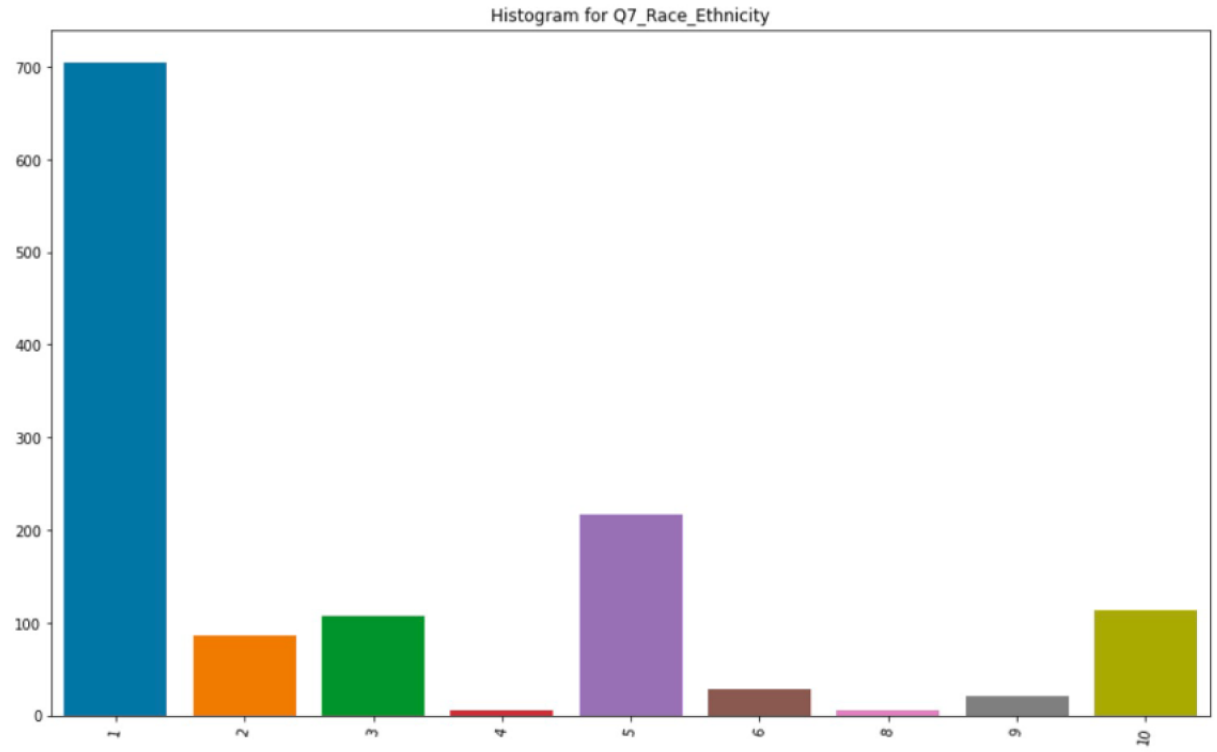
Observation of industry differences for Wrigley dataframe:
1. We find that for Petcare, the 20s age group(4) has the largest proportion, the second one is group(3), but number of group(5) is close to group(3)
2. For the gender column of Wrigley, there are small portion of "Hyper-Masculine(1)", the number of "Masculine(4)" and "Feminine(2)" is close, which is same as sex column
3. From barplot we find that for Confectionary segment, the amount of male is 50.4% which is more than female 49.6%
4. By observing barplot, the number of "males(2)" and "females(1)" is basically the same
5. From the four barplot for sex portions of different industry, we can conclude that the Wrigley has the smallest difference in the number of male and female
6. We get the average value of Q2_Age for the Wrigley industry is 4.29(around 25 years old), which is smaller than Confectionary and Petcare and bigger than Food, the std value is 1.15, means the age distribution is relatively concentrated



Observation of industry differences for Food dataframe:
1. We find that for Petcare, the 20s age group(4) has the largest proportion, the second one is group(5), and the number of group(1) is close to group(5)
2. For the gender column of Wrigley, there are large portion of "Not Applicable(10)", the number of "Masculine(4)" and "Feminine(2)" is close, which is same as sex column
3. From barplot we find that for Confectionary segment, the amount of "Male(2)" is 50.5% which is more than "Female(1)" 49.5%
4. By observing barplot, the number of "Male(2)" and "Female(1)" is basically the same
5. We get the average value of Q2_Age for the Food industry is 3.99(around 20 years old), which is smallest among these four industry, the std value is 1.98, means the age distribution is not relatively concentrated
6. We can conclude that the the average age of the Food industry is the youngest



For the Q7_Race_Ethnicity column, we find that value "1"("White") is the largest, the second one and third one are "5"("Asian/Asian American") and "10"("Not Applicable"). We can find that the number of whites is greater than the sum of the numbers of all other races.



Industry differences of the Q7_Race_Ethnicity column

We can find that among the Q7_Race_Ethnicity = 1("White"), the number of "Confectionary" is 310 and "Petcare " is 263 which are far more than "Wrigley" and "Food" industry. Majority of "White" work in these two industries.

We can find that among the Q7_Race_Ethnicity = 2("Hispanic/Latino"), the "Confectionary" is much higher than other industries, each industry has "Hispanic/Latino", most of "Hispanic" are in "Confectionary" industry. We can find that among the Q7_Race_Ethnicity = 3("Black"), the "Confectionary" is also higher than other industries, each industry has "Black". The portion of "Petcare" for "Black" is the highest compared with other ethnicities.

We can find that among the Q7_Race_Ethnicity = 4(Native American/Hawaiian/Alaskan/Pacific Islander), there is a small amount of data. We can find that among the Q7_Race_Ethnicity = 5("Asian/Asian American"), there are a very high number of people working in the "Confectionary" industry.

We can find that among the Q7_Race_Ethnicity = 6("Middle Eastern"), no one is in the "Food" industry. The ratio of "Confectionary" among industries for "Asian" is the highest compared with other ethnicities.

We can find that among the Q7_Race_Ethnicity = 8("Mixed Race"), there are small amount of data, the number of "Confectionary" is 3, for "Petcare" is 2, for "Wrigley" is 1, no one is in "Food" industry.

We can find that among the Q7_Race_Ethnicity = 9("Can't tell"), the number of data is relatively small, the number of people for "Confectionary" and "Food" is the same.

We can find that among the Q7_Race_Ethnicity = 10("Not Applicable"), the "Confectionary" is higher than other industries, people are only in the "Confectionary" and "Food" industry.

By the statistics, we observe that the largest amount of data whatever their race/ethnicity are in the "Confectionary" industry. And many of people are "Not Applicable" of Q7_Race_Ethnicity in "Confectionary" and "Food".

Result for Question 2: Is there a change in representation in advertisements over time?

First we have an observation for the entire dataset:

1. For Year_Aired column, there are only 2 unique values: 2019 and 2020.

2. For Q2_Age column, most of values are Q2_Age = 4, which is the age of 20-29 year olds

3. For Q3_Sex column, number of male is more than female, and there is a little people don't tell their sex

4. For Q7_Race_Ethnicity column, we find that value 1 is the largest, which means white people.

Since "time" refers to the "Year_Aired" column, and "representation" refers to all the question columns. We find the differences among 'Q2_Age', 'Q3_Sex' and 'Q7_Race_Ethnicity' columns over Year_Aired, since they are common statistical variables

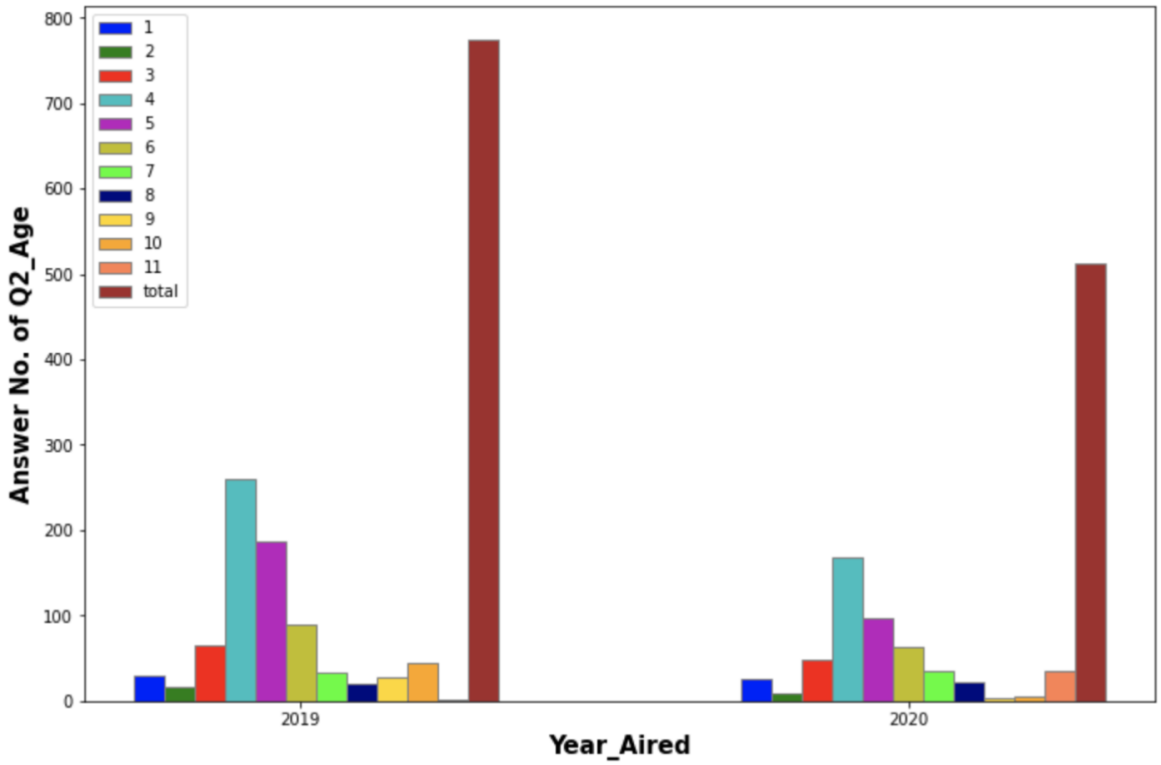
Observation of differences for Q2_Age over time:

1. The total number decreases from 2019 to 2020

2. Q2_Age = 4 is the largest value between two years

3. Someone don't tell or are unwilling to disclose the information of age

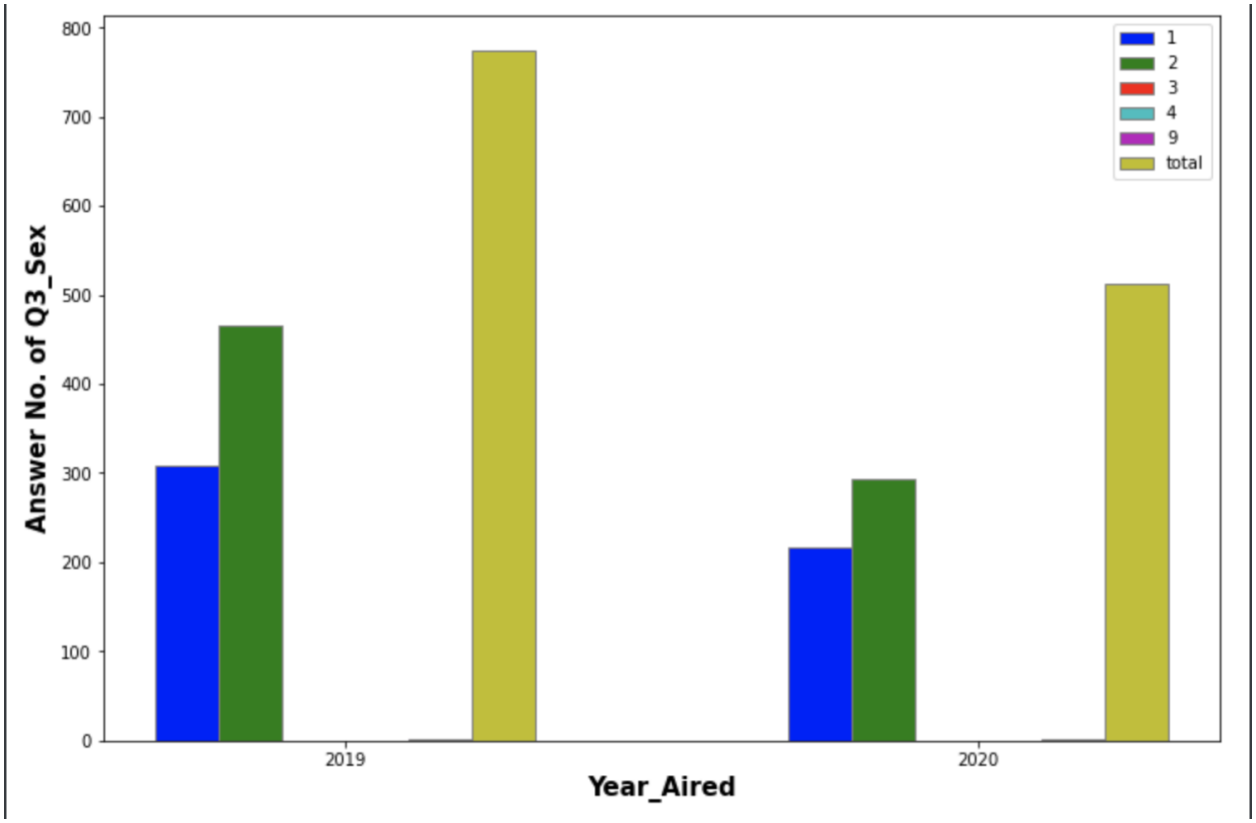
4. The distributions of ages between two years are almost the same.



Observation of industry differences for Q3_Sex:

1. We find that the total number decreases from 2019 to 2020

2. Male numbers are larger than female numbers in both two years 3. And some of people can not tell their sex

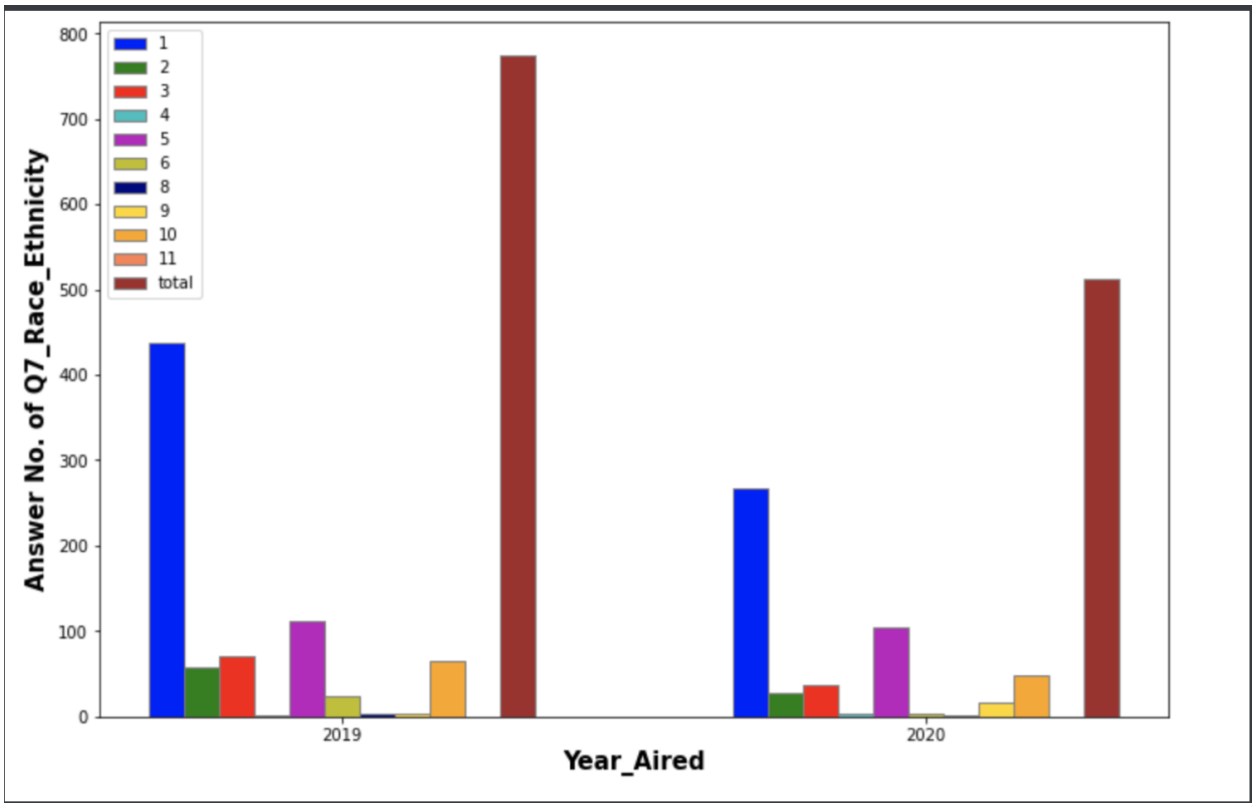


Observation of industry differences for Q7_Race_Ethnicity:

1. We find that the total number of Q7_Race_Ethnicity decreases from 2019 to 2020

2. Among all different answers, option 1, which refers to "white people", is the largest in both two years

3. The total number of rest options is almost the same as the number of option 1.



Result for Question 3: What are the trends?

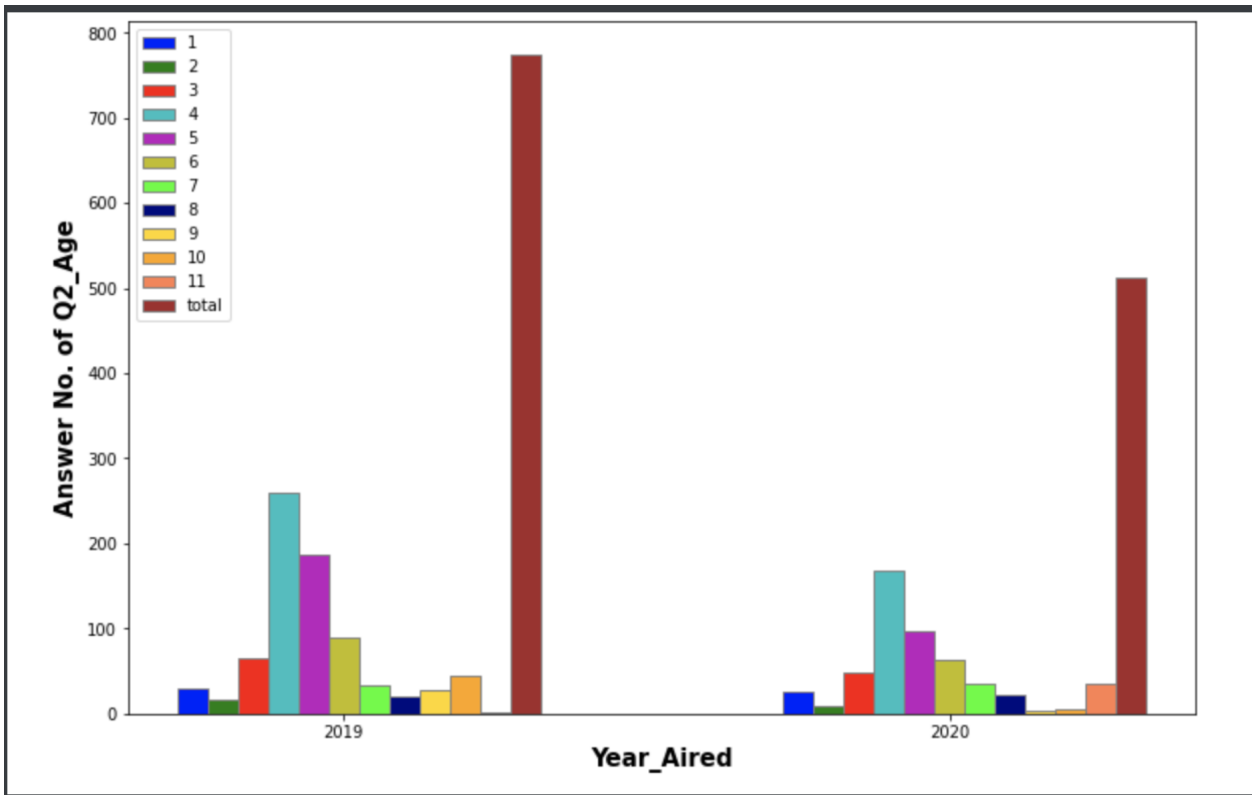
Our group tackled this question based on our previous answer to question "Is there a change in representation in advertisements over time?" Our interpretation of "trends" is: the data moving tendency in terms of the columns (e.g. age, sex) over time (i.e. year_aired).

Trends of Q2_Age:

1. The total number decreases from 2019 to 2020

2. Q2_Age = 4 is the largest value between two years

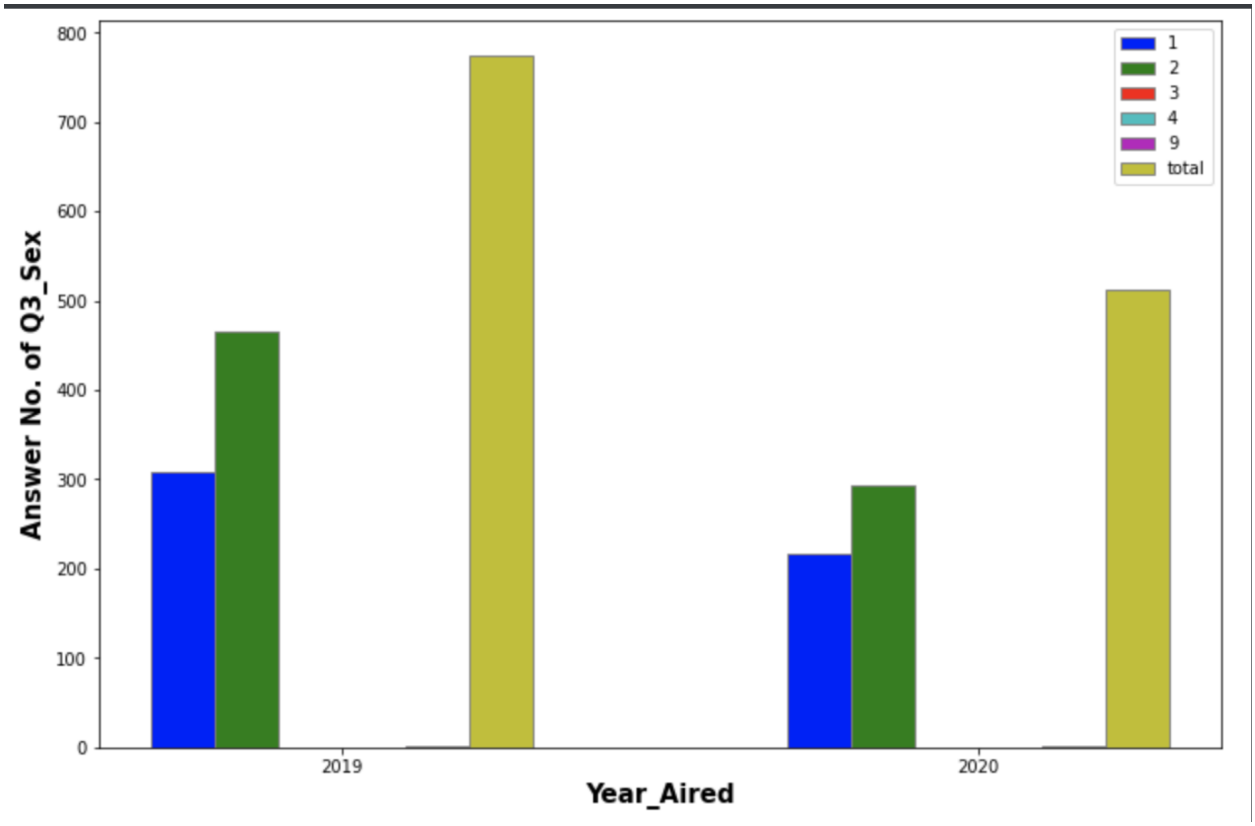
3. Someone doesn't tell or are unwilling to disclose the information of age 4. The distributions of ages between two years are almost the same.



Trends of Q3_Sex:

1. We find that the total number decreases from 2019 to 2020

2. Male numbers are larger than female numbers in both two years 3. And some of people can not tell their sex

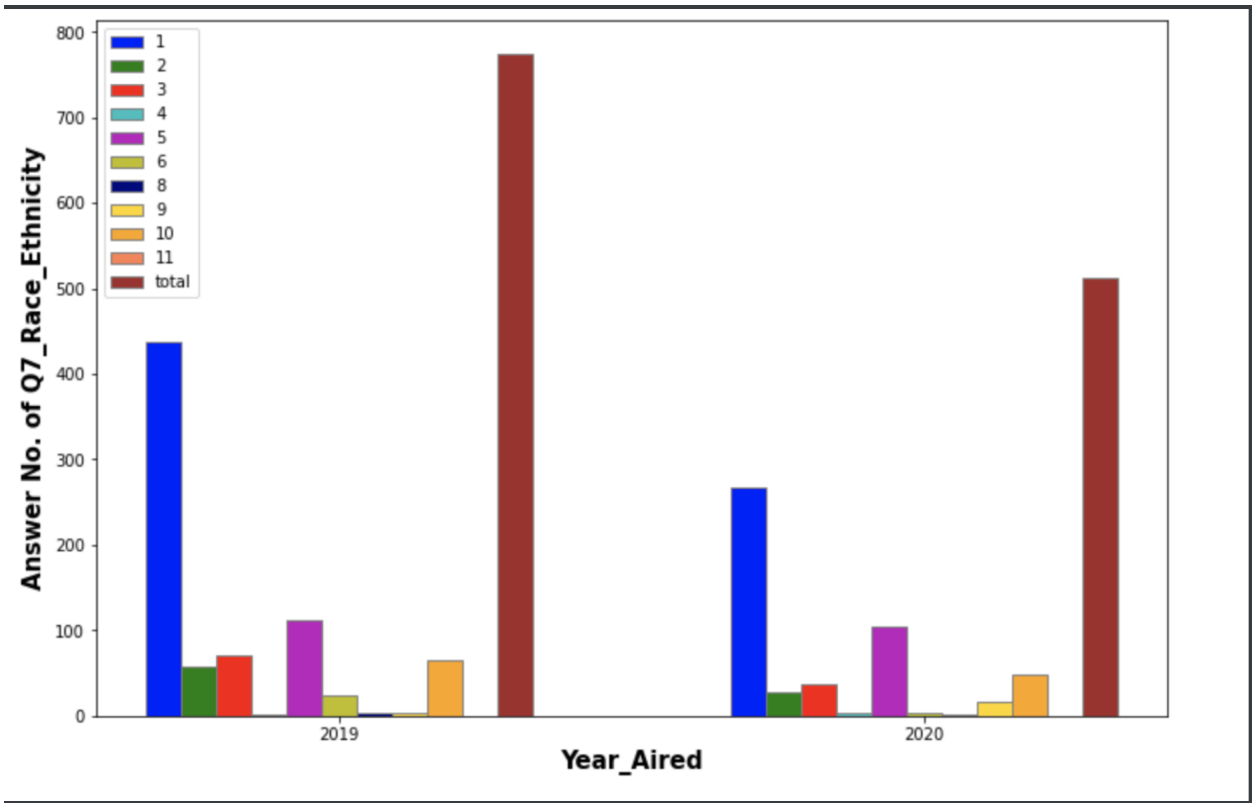


Trends of Q7_Race_Ethnicity:

1. We find that the total number of Q7_Race_Ethnicity decreases from 2019 to 2020

2. Among all different answers, option 1, which refers to "white people", is the largest in both two years

3. The total number of rest options is almost the same as the number of option 1.



Appendix

This section will demonstrate all the code that our group has written to normalize, merge, and analyze the datasets.

Q1:

In [1]:
```
import numpy as np
import pandas as pd
import warnings
warnings.filterwarnings("ignore")
import matplotlib.pyplot as plt
import seaborn as sns
```

Analysis whole data set

In [2]:
```
df = pd.read_csv("Merged_Preprocessed_Mars2020_2021.csv")
```

Show all columns

In [3]:
```
df.columns
```

Out[3]:
```
Index(['Coder', 'Asset_Name', 'Brand', 'Lead_Country', 'Year_Produced',
       'Year_Aired', 'Segment', 'Agency', 'Character_Name',
       'Character_Description', 'Q1_Prominence', 'Q2_Age',
       'Q2a_Age_SpecifyOther', 'Q3_Sex', 'Q4_Gender', 'Q5_LGBTQ',
       'Q6_Animated', 'Q6a_Animated_OtherSpecify', 'Q7_Race_Ethnicity',
       'Q7a_RE_OtherSpecify', 'Q8_Disability', 'Q9a_Shopping', 'Q9b_Driving',
       'Q9c_Cleaning', 'Q9d_Cooking', 'Q9e_Working', 'Q9f_Socializing',
       'Q9g_Nothing', 'Q9h_Eating/Drinking', 'Q9i_Exercising',
       'Q9j_Activity Other', 'Q9j_Activity_Other_Specify', 'Q10a_Kitchen',
       'Q10b_Office', 'Q10c_Car', 'Q10d_Store', 'Q10e_Outdoors',
       'Q10f_Living Room', 'Q10g_Restaurant_Bar', 'Q10h_Gym', 'Q10i_Bedroom',
       'Q10j_Bathroom', 'Q10k_Sporting_Event', 'Q10l_Classroom',
       'Q10m_Location_Other', 'Q10m_Location_Other_Specify',
       'Q11_Revealing_Clothing', 'Q12_Nudity', 'Q12a_Nudity_Other_Specify',
       'Q13_Visually_Objectified', 'Q14_Verbally_Objectified',
       'Q15_Intelligence', 'Q16_Humor', 'Q17_Occupation', 'Q18_Leader',
       'Q19_Authority', 'Q20_Body Type', 'Q20a_Body_Type_Other_Specify',
       'Q21a_Disordered_Eating', 'Q21b_Selfy_injury', 'Q21c_NegativeTalk',
       'Q21d_Body_Modification', 'Q22a_Visual_Shame', 'Q22b_Verbal_Shame',
       'Q22c_Sizeist_Slurs', 'Q22d_Punchline',
       'Q22e_Denied_Personal_Opportunity',
       'Q22f_Denied_Professional_Opportunity', 'Q22g_Other_Prejudice',
       'Q22g_Prejudice_Other_Specify', 'Q23a_Lazy', 'Q23b_Physically_Slow',
       'Q23c_Stupid', 'Q23d_Loser', 'Q23e_Inactive', 'Q23f_Poorly_Dressed',
       'Q23g_Funny', 'Q23h_Jolly', 'Q23i_Clumsy', 'Q23j_Alone',
       'Q24a_Comic_Relief', 'Q24b_Sidekick', 'Q24c_Mamma_Hen', 'Q24d_Nympho',
       'Q25_Fat_to_Fit', 'Q26_Inspo_Porn', 'Q25_NOTES', 'Notes on Dwelling'],
      dtype='object')
```

Seperate to four dataframe of different columns

In [4]:
```
segment_col = df[['Segment']]

age_col = df[['Q2_Age','Segment']]

sex_col = df[['Q3_Sex','Segment']]

gender_col = df[['Q4_Gender','Segment']]
```

category feature nunique distribution

In [5]:
```
# classify feature
cat_fea =['Segment','Q2_Age','Q3_Sex','Q4_Gender']
# category feature nunique distribution
for fea in cat_fea:
    print('*****************************')
    print(fea + "'s feature distribution as follow': ")
    print("{} feature has {} different values".format(fea,df[fea].nunique()))
    print(df[fea].value_counts())
```

```
*****************************
Segment's feature distribution as follow':
Segment feature has 4 different values
Confectionary    729
Petcare          346
Wrigley          119
Food              93
Name: Segment, dtype: int64
*****************************
Q2_Age's feature distribution as follow':
Q2_Age feature has 11 different values
4     428
5     283
6     152
3     115
7      68
1      56
10     50
8      43
11     36
9      31
2      25
Name: Q2_Age, dtype: int64
*****************************
Q3_Sex's feature distribution as follow':
Q3_Sex feature has 3 different values
2    760
1    524
9      3
Name: Q3_Sex, dtype: int64
*****************************
Q4_Gender's feature distribution as follow':
Q4_Gender feature has 5 different values
10    512
2     450
4     318
1       5
9       2
Name: Q4_Gender, dtype: int64
```

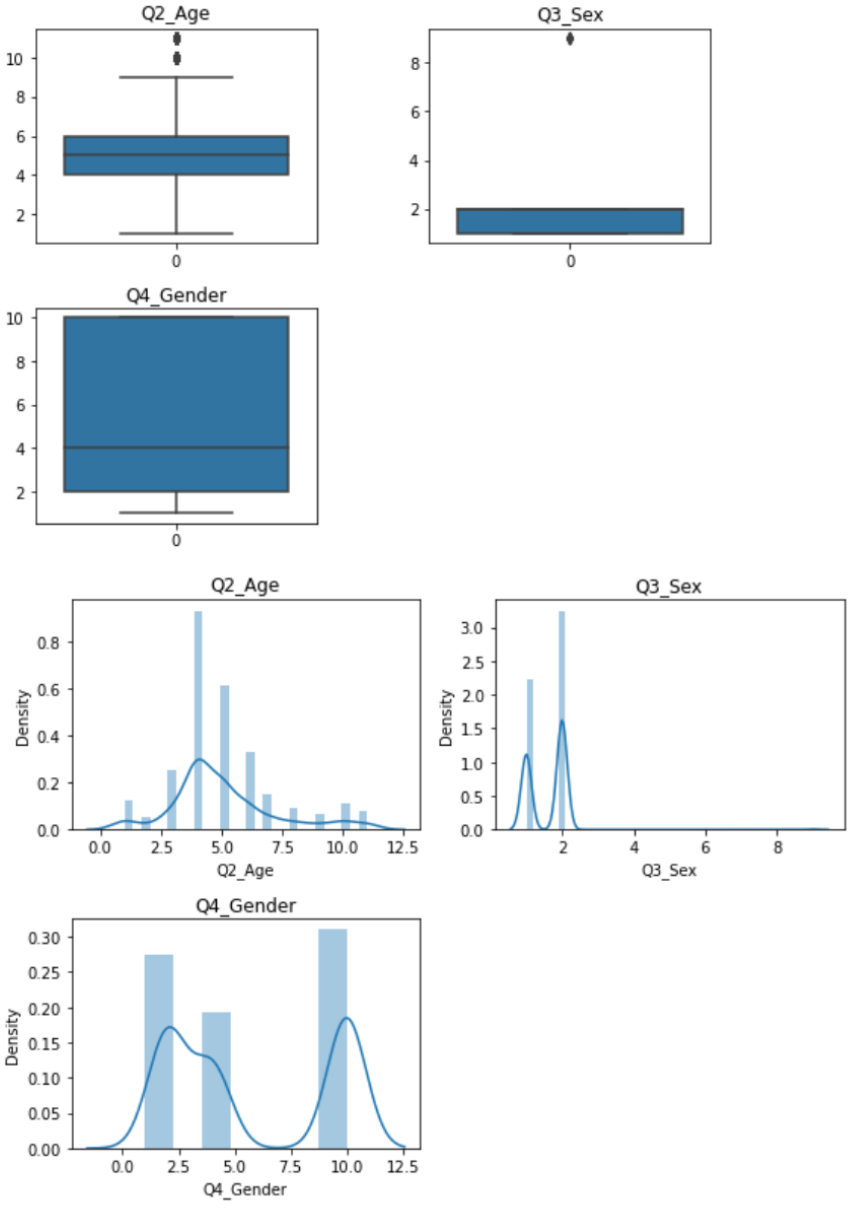
Plot a boxplot and linegraph of numeric data and check/detect for outliers

In [6]:
```
fig=plt.figure(figsize=(8,6))
i=1
num_fea =['Q2_Age','Q3_Sex','Q4_Gender']
for col in num_fea:
    ax=fig.add_subplot(2,2,i)
    sns.boxplot(data=df[col],ax=ax)
    i=i+1
    plt.title(col)

plt.subplots_adjust(wspace=0.4,hspace=0.3)
plt.show()

fig=plt.figure(figsize=(8,6))
i=1
for col in num_fea:
    ax=fig.add_subplot(2,2,i)
    sns.distplot(df[col])
    i=i+1
    plt.title(col)

plt.tight_layout()
```



In [7]:
```
# plot histogram to show 10 most
plt.figure()
plt.figure(figsize=(24, 24))
i = 1
for fea in cat_fea:
        plt.subplot(3, 2, i)

        i += 1
        v = df[fea].value_counts().nlargest(15)
        fig = sns.barplot(x=v.index, y=v.values)
        for item in fig.get_xticklabels():
            item.set_rotation(80)
        plt.title('Histogram for '+fea)
plt.tight_layout()
plt.show()
```

Question :

Are there industry differences in representation in advertisements?

“industry” refers to the “Segment” column, and “representation” refers to all the question columns.

In [8]:
```
industry = segment_col
```

In [9]:
```
industry.describe()
```

Out[9]:

	Segment
count	1287
unique	4
top	Confectionary
freq	729

Specific examples of how to interpret the questions could be: Are there segment differences in different ages?

Industry differences of the Q2_Age column

In [10]:
```
age_col.head()
```

Out[10]:

	Q2_Age	Segment
0	5	Confectionary
1	4	Confectionary
2	8	Confectionary
3	2	Confectionary
4	2	Confectionary

In [11]:
```
for i in range(1,12):
    df_age = age_col.loc[age_col['Q2_Age'] == i]
    print('Q2_Age =',i)
    print(df_age['Segment'].value_counts())
    print()
```

Observation for Q2_Age:

We can find that among the Q2_Age=1, the Petcare is the most

With the increasing of the age, more people are in the "Confectionary" industry, but when Q2_Age = 8 the Petcare is greater than Confectionary

And by the statistics, we observe that the largest amount of data are in the "Confectionary" industry

And someone don't tell or are unwilling to disclose the information of age

Industry differences of the Q3_Sex column

In [12]:
```
sex_list = [1,2,9]
for i in sex_list:
    df_sex = sex_col.loc[sex_col['Q3_Sex'] == i]
    print('Q3_Sex =',i)
    print(df_sex['Segment'].value_counts())
    print()
```

```
Q3_Sex = 1
Confectionary    262
Petcare          157
Wrigley           59
Food              46
Name: Segment, dtype: int64

Q3_Sex = 2
Confectionary    467
Petcare          186
Wrigley           60
Food              47
Name: Segment, dtype: int64

Q3_Sex = 9
Petcare    3
Name: Segment, dtype: int64
```

Observation for Q3_Sex:

We can find that among the Q3_Sex = 1 which is "Female", the "Confectionary" is the most, "Food" is the smallest

We can find that among the Q3_Sex = 2 which is "Male", the "Confectionary" is the most, "Food" is the smallest

We can find that among the Q3_Sex = 9 which is "Can't tell", there is only Petcare

And by the statistics, we observe that the largest amount of data are in the "Confectionary" industry

And some of people can not tell their sex

Industry differences of the Q4_Gender column

In [13]:
```
gender_list = [1,2,4,9,10,]
for i in gender_list:
    df_gender = gender_col.loc[gender_col['Q4_Gender'] == i]
    print('Q4_Gender =',i)
    print(df_gender['Segment'].value_counts())
    print()
```

Observation for Q4_Gender:

We can find that among the Q4_Gender = 1 which is "Hyper-Masculine", is in either Petcare or Wrigley

We can find that among the Q4_Gender = 2 which is "Masculine", the "Confectionary" is much higher than other industries, each industry has "Masculine"

We can find that among the Q4_Gender = 4 which is "Feminine", the "Confectionary" is also higher than other industries, each industry has "Feminine"

And by the statistics, we observe that the largest amount of data whatever their gender are in the "Confectionary" industry

And many of people are "Not Applicable" of Q4_gender in "Confectionary", "Petcare" and "Food"

Try different way to answer the question, separate four different segments to different dataframes

Create confectionary dataframe

In [14]:
```
df_confectionary = df.loc[df['Segment'] == 'Confectionary']
df_confectionary = df_confectionary.loc[:,['Q2_Age','Q3_Sex', 'Q4_Gender']]
df_confectionary.reset_index(inplace = True)
df_confectionary.drop('index', axis = 1, inplace = True)
df_confectionary.head()
```

Out[14]:

	Q2_Age	Q3_Sex	Q4_Gender
0	5	2	10
1	4	2	10
2	8	2	10
3	2	2	10
4	2	1	10

category feature nunique distribution

In [15]:
```
cat_fea_confectionary =['Q2_Age','Q3_Sex','Q4_Gender']
# category feature nunique distribution
for fea in cat_fea_confectionary:
    print('*****************************')
    print(fea + "'s feature distribution as follow': ")
    print("{} feature has {} different values".format(fea,df_confectionary[fea].nunique()))
    print(df_confectionary[fea].value_counts())
```

```
*****************************
Q2_Age's feature distribution as follow':
Q2_Age feature has 11 different values
4     246
5     134
6      85
3      75
10     48
7      38
11     31
9      28
8      17
2      14
1      13
Name: Q2_Age, dtype: int64
*****************************
Q3_Sex's feature distribution as follow':
Q3_Sex feature has 2 different values
2    467
1    262
Name: Q3_Sex, dtype: int64
*****************************
Q4_Gender's feature distribution as follow':
Q4_Gender feature has 4 different values
2     313
10    264
4     151
9       1
Name: Q4_Gender, dtype: int64
```

Plot Histogram for columns

In [16]:
```
# plot histogram to show
plt.figure()
plt.figure(figsize=(24, 24))
i = 1
for fea in cat_fea_confectionary:
        plt.subplot(3, 2, i)
        i += 1
        v = df_confectionary[fea].value_counts().nlargest(15)
        fig = sns.barplot(x = v.index, y = v.values)
        for item in fig.get_xticklabels():
            item.set_rotation(80)
        plt.title('10 Most '+fea)
plt.tight_layout()
plt.show()
```

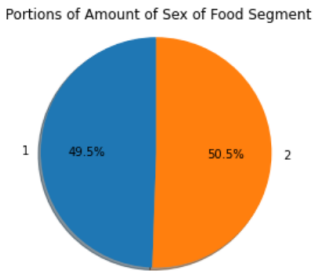
```
<Figure size 432x288 with 0 Axes>
```

Barplot

In [32]:
```
Sex1_food = df_food[df_food['Q3_Sex'] == 1]['Q3_Sex'].value_counts()
Sex2_food = df_food[df_food['Q3_Sex'] == 2]['Q3_Sex'].value_counts()

labels_food = '1','2'
sizes_food = [Sex1_food.values.item(), Sex2_food.values.item()]
explode_food = (0, 0)  # only "explode" the 2nd slice (i.e. 'Hogs')

fig, ax = plt.subplots()
ax.pie(sizes_food, explode = explode_food, labels = labels_food, autopct='%1.1f%%',
        shadow = True, startangle = 90)
ax.axis('equal')  # Equal aspect ratio ensures that pie is drawn as a circle.
plt.title(' Portions of Amount of Sex of Food Segment')
plt.show()
```



We can find that for Confectionary segment, the amount of "Male(2)" is 50.5% which is more than "Female(1)" 49.5%

The number of "Male(2)" and "Female(1)" is basically the same

Compute Mean and std of Q3_Age for Petcare

In [33]:
```
age_mean_food = df_food['Q2_Age'].mean()
age_mean_food = round(age_mean_food, 2)
age_std_food = df_food['Q2_Age'].std()
age_std_food = round(age_std_food, 2)
print('Mean of Age is', age_mean_food, 'Standard deviation of Age is',age_std_food)
```