Final Project Draft

18/04/2022

—

Sponsors
Meredith Conroy

Romeo Perez

Project Manager
Grace Yoon

Team Members
Dayong Wu (Team Leader)

Emmanouil Kritharakis

Yan Tong

Junfei Huang

Yuanli Wang

Background

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

After exploring the datasets, we can then do formal analysis. We performed data analysis mainly based on the key questions listed in the project proposal.

- We can find that among the Ethnicity column, value = "White", the number of "Confectionary" industry and "Petcare " industry which are far more than "Wrigley" and "Food" industry. Therefore, base on our dataset, the majority of "White" work in "Confectionary" and "Petcare" industries.
- We observe that the largest number of people work in "Confectionary" industry regardless of gender.
- For Age, most of value is 20-29 years old.

Histogram of "Confectionary" for Q2_Age
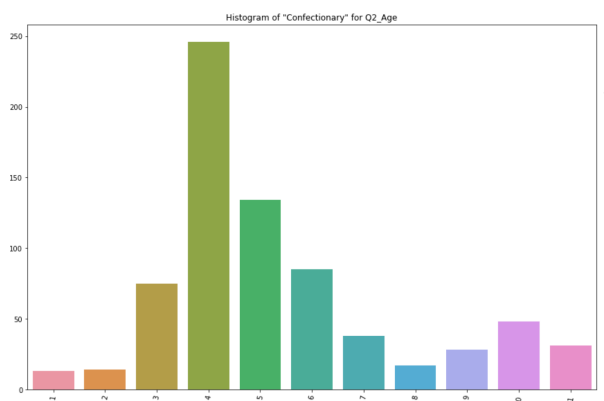
Race Portions of Amount of Confectionary Industry