



CS-401

# Applied Data Analysis

## P1 Milestone

### Authors

Amélie Menoud

András Horkay

Cyrielle Manissadjian

Dominic Stratila

Katia Todorov

## Andras's Ideas

I am going to put my submission here so that when we compare and pick the project we will do, we can compare them easily

### When I Grow Up, I'll Be Anything but a Politician: Professions and Social Attitudes in the New Yorker Caption Contest

This project explores how different professions are portrayed in the New Yorker Caption Contest and how audiences respond to them. While some jobs, such as doctors, are often admired, others, like politicians and lawyers, are frequent targets of humour. Our research will focus on, but might not be limited to: Which professions appear most often in captions? Are certain occupations consistently rated funnier? Are some portrayed in a positive or negative light? To address these questions, we will first identify professions mentioned in captions using a U.S. Census occupation list, accounting also for indirect workplace references (e.g., "courtroom" for lawyers). The professions will then be grouped into broader categories (healthcare, law, education, politics, etc.). Finally, we will analyse their frequency, average funniness scores, and recurring stereotypes. We hope the findings will shed light on societal attitudes towards various professions as reflected in popular culture, and perhaps give ideas which professions to avoid in future careers to avoid being the butt of jokes.

### Laughing at the Forbidden: Taboo Themes in the New Yorker Caption Contest

This project investigates how taboo themes appear in the New Yorker Caption Contest and how audiences respond to them. While taboo jokes may be funny due to shock value, they can also be offensive or harmful. Some key questions we will answer are: Which taboo topics occur most often? How are they rated in terms of funniness? Do their frequencies change over time or with the political landscape in the US? To answer these, we will define taboo categories (e.g., suicide/self-harm, racism, sexism, violence, religion) and identify their presence in captions through keywords and phrases. We will then analyse the frequency at which these topics occur, their associated funniness scores, and any temporal trends, paying attention to events such as elections or economic crises that may influence prevalence. The project will shed light on how societal attitudes towards taboo topics are reflected in humour and how audience perceptions of these themes evolve over time.

### Funny or Stereotypical? Gender Roles in the New Yorker Caption Contest

This project examines how gender is represented in the New Yorker Caption Contest, both in cartoons and in audience-submitted captions. The key questions we will attempt to answer are: Do men appear more frequently than women in the images? When women are depicted, are they more often shown in stereotypical roles (domestic or caregiving) rather than professional settings (as leaders)? In captions, are gendered terms or phrases used that reinforce stereotypes? We will analyse cartoon metadata to measure representation, classify roles to detect stereotypes, and examine captions for gendered language. These steps may involve image as well as text analysis. Finally, we will compare funniness scores to assess whether captions using gendered language or stereotypes are rated differently by audiences. The results will provide insight into the persistence of gender stereotypes in humour and their reception by the public.