

Natural Language Processing on

TED Talks

Ideas worth spreading

Motivation

- **TED Talks**
 - Online Influential videos
- **Leverage NLP capabilities**
 - Topic Modeling
 - Sentiment Analysis
- **Predict views**
 - Linear regression

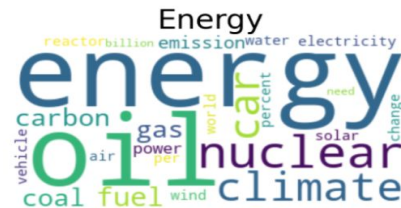
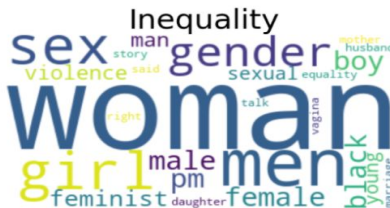
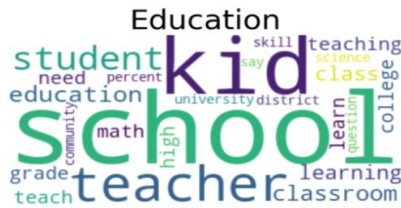
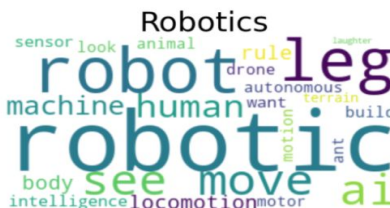
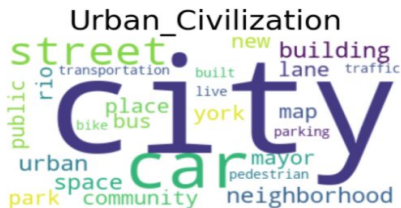


Data

- **Kaggle**
 - 2467 talk transcripts
 - 2006 to 2017
- **Preprocessing**
 - Regex
 - NLTK
- **Topic Modeling**
 - Tfidf and NMF

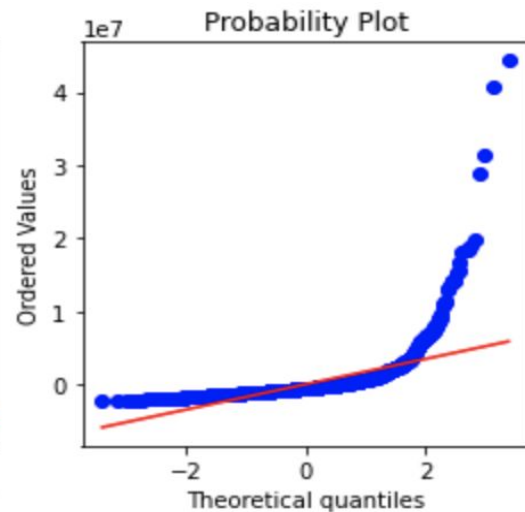
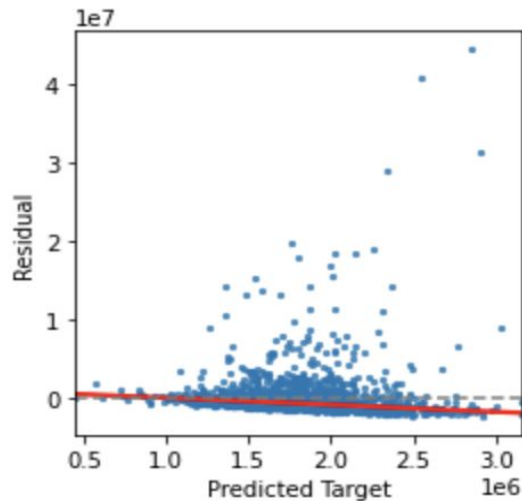
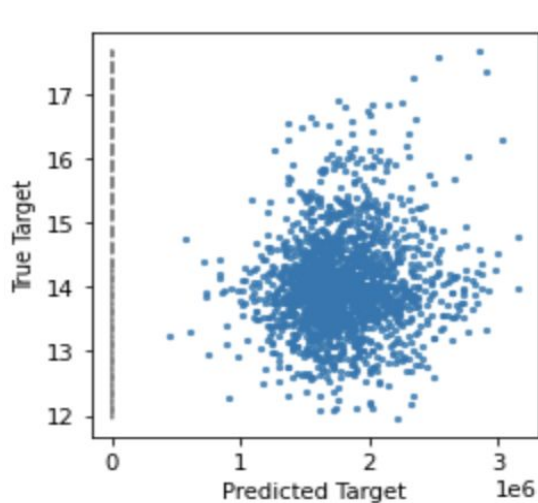


Topic Modelling



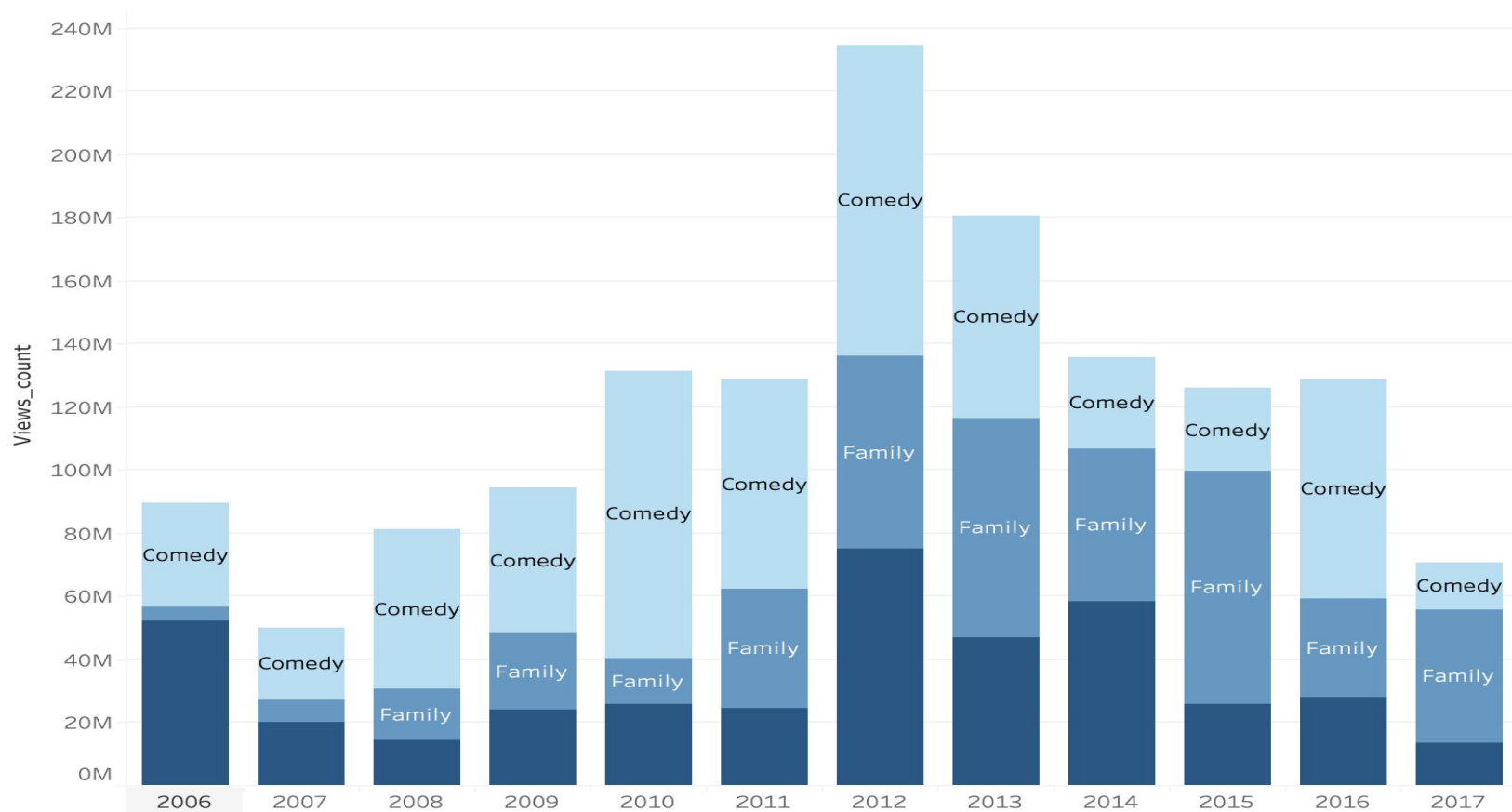
Linear Regression with log transformation

Diagnostic Plots

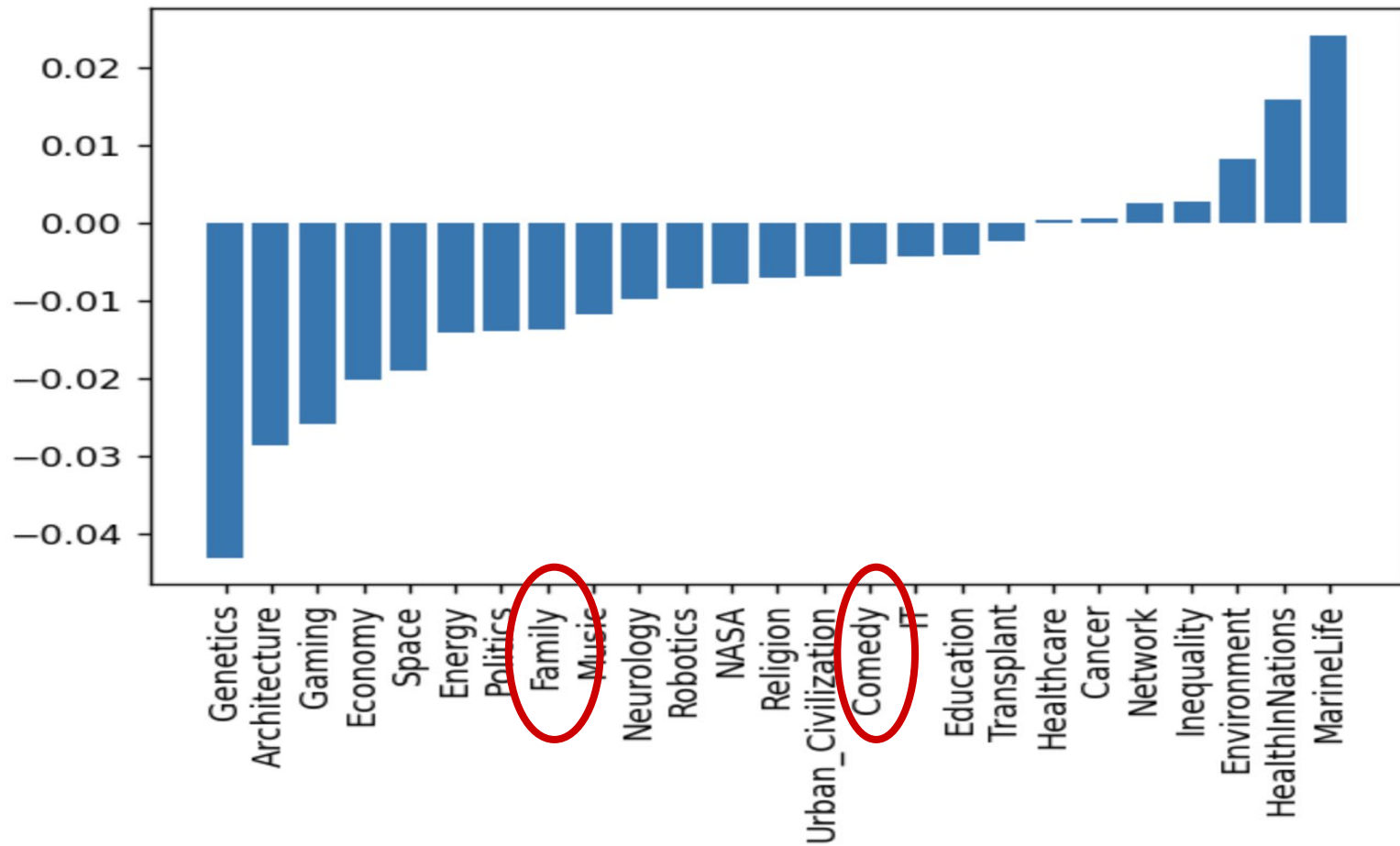


R2: 0.003
RMSE: 2581606

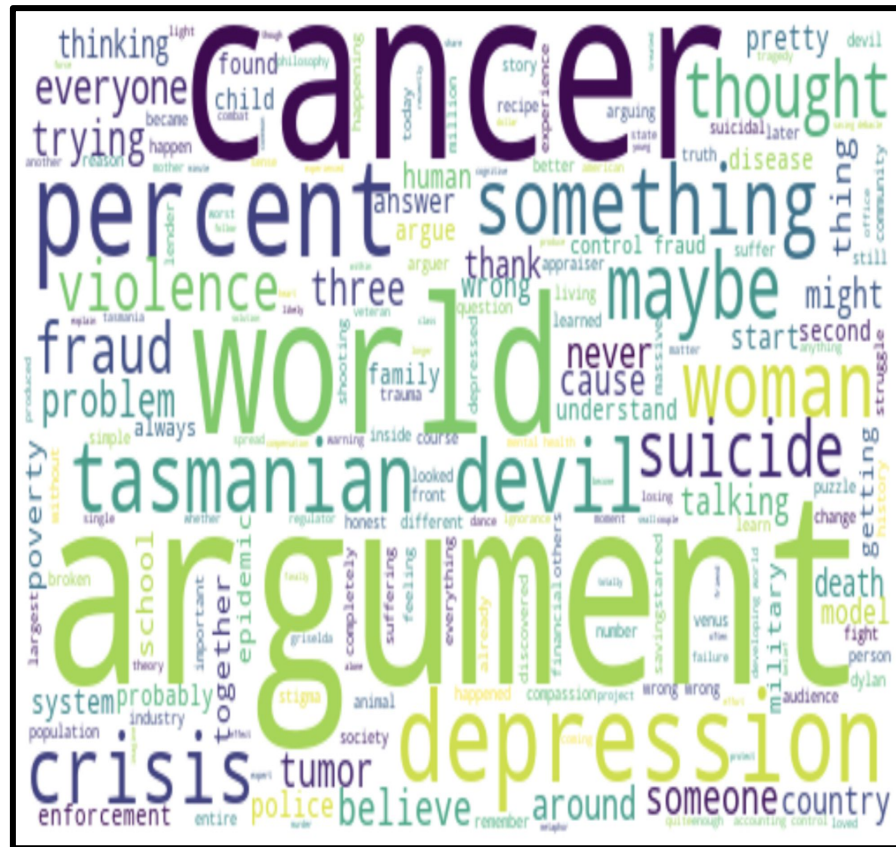
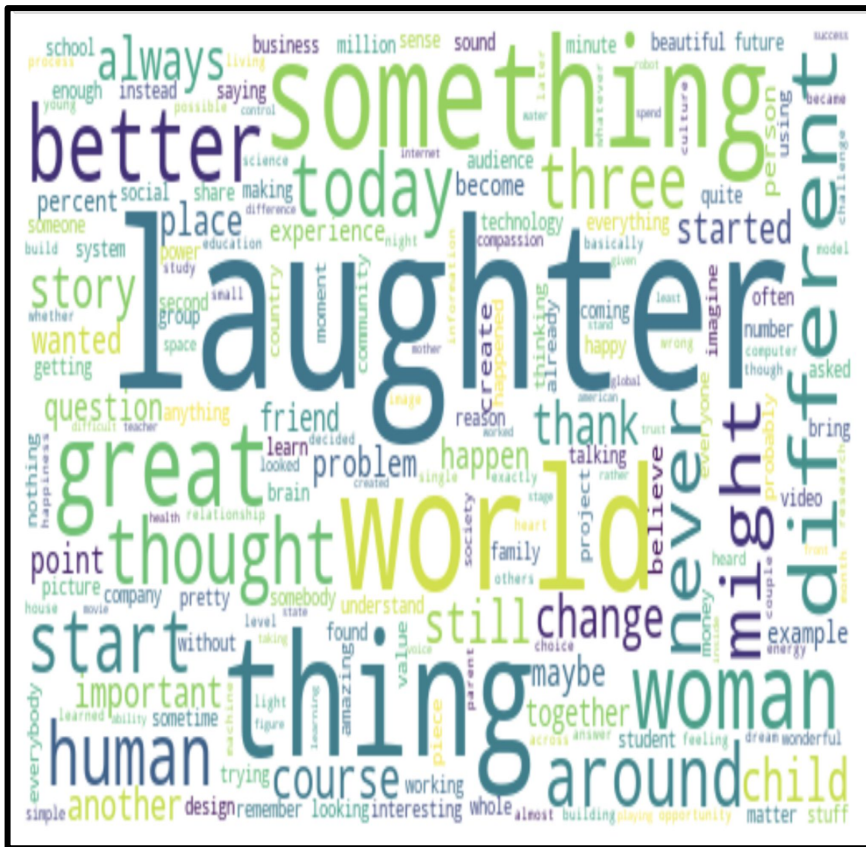
Top Topics



Feature Coefficients



Sentiment Analysis



Future Work

- Try other regression models
- More NLP techniques
- Recommendation Engine

Thank you!