Towards the Perfect Blog Post

Bryan Ross

towards
data science

Why TDS?

- COMMUNITY

- EXPOSURE

- MEDIUM PARTNER PROGRAM

SCRAPE ➡ PROCESS ➡ MODEL ➡ EXTRACT

BeautifulSoup
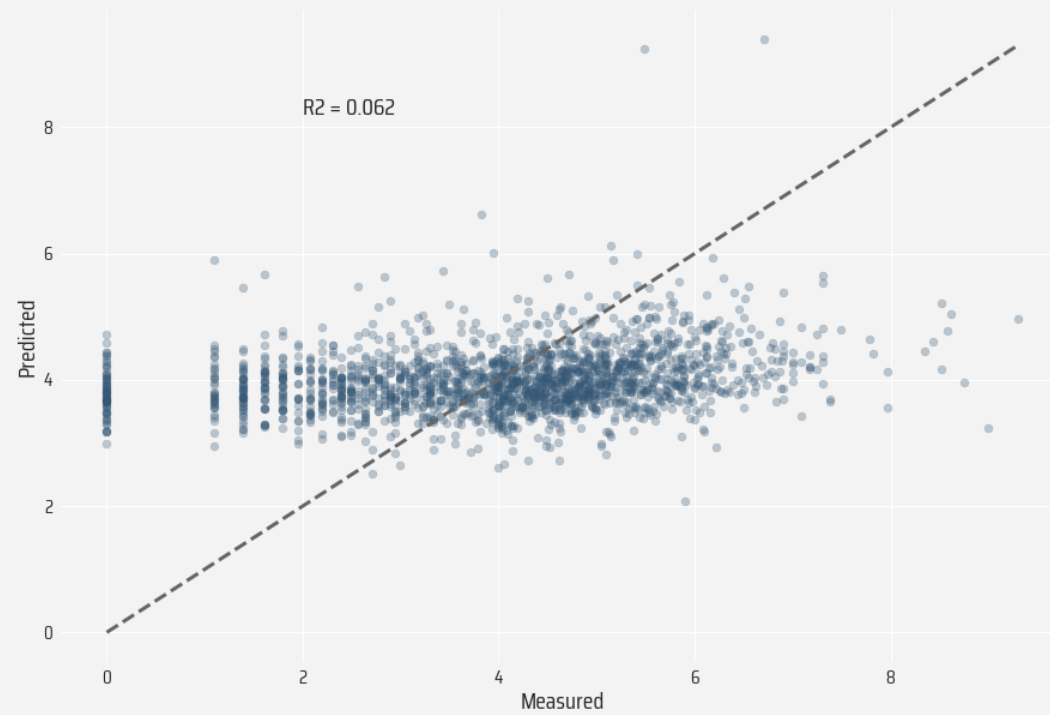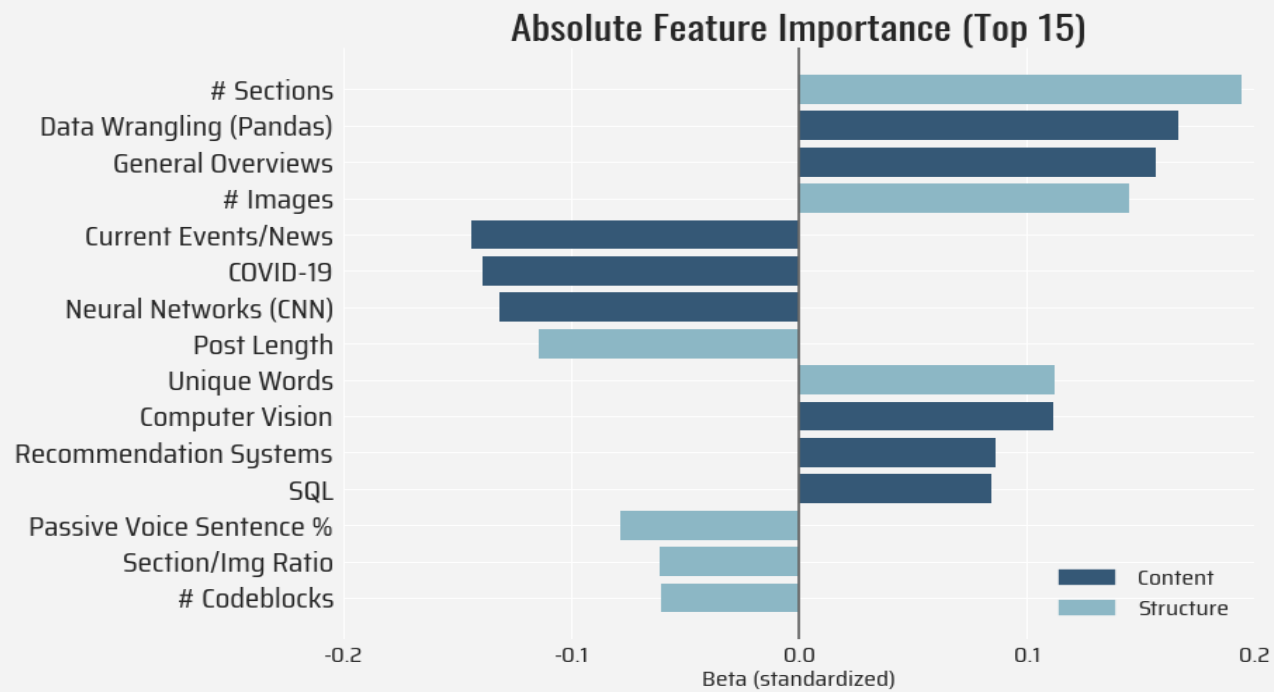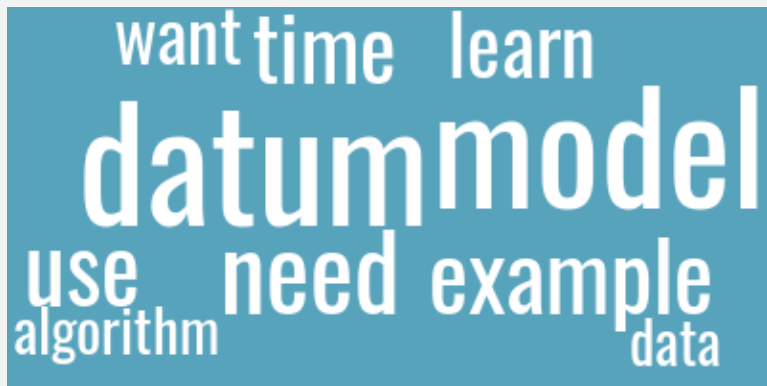
spaCy

statsmodels

Se

pandas

NumPy

scikit learn

Prediction Error Plot

BOTH CONTENT AND STRUCTURE MATTER

Absolute Feature Importance (Top 15)

Legend: Content, Structure

X-axis: Beta (standardized)
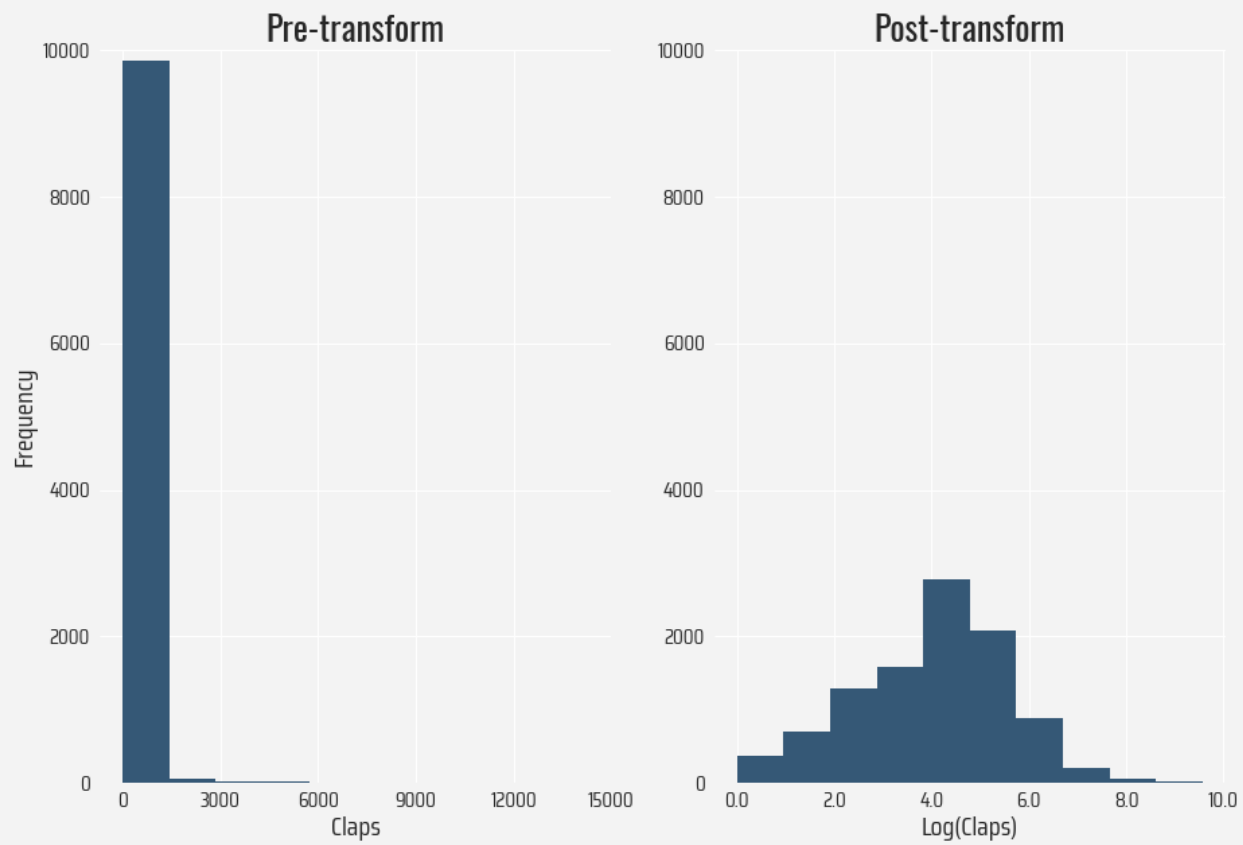
General Overviews

COVID-19

Takeaways

- Avoid long blocks of text, include visuals

- Use active voice, vary vocabulary
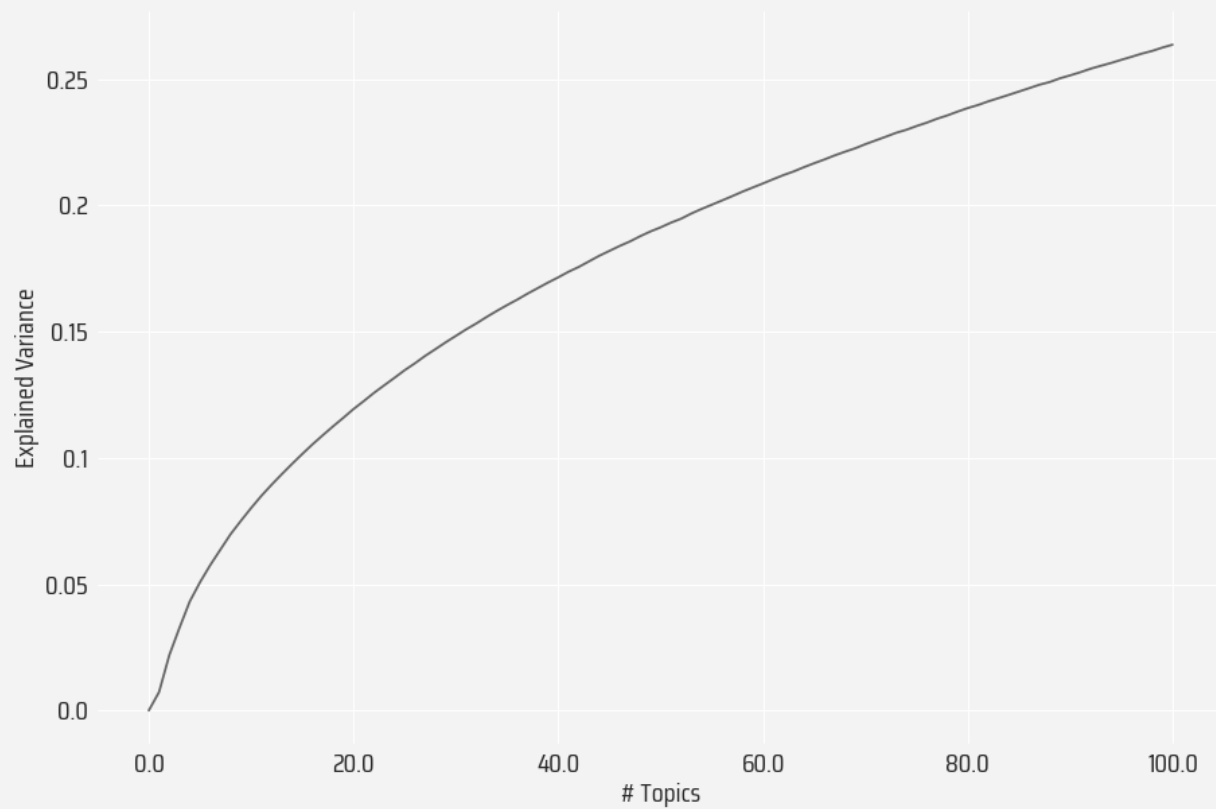
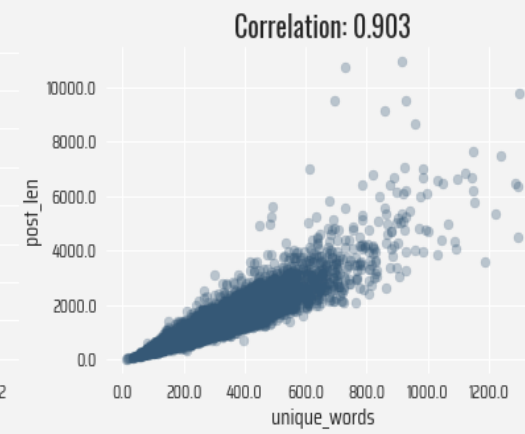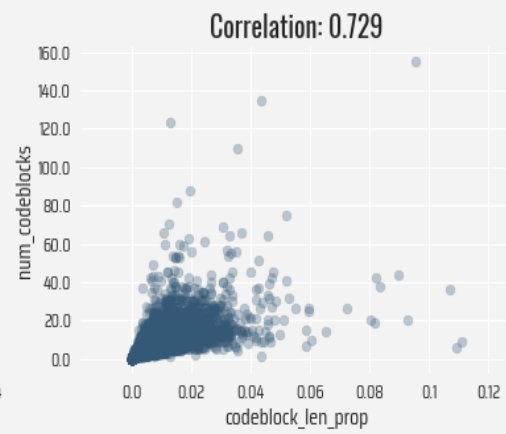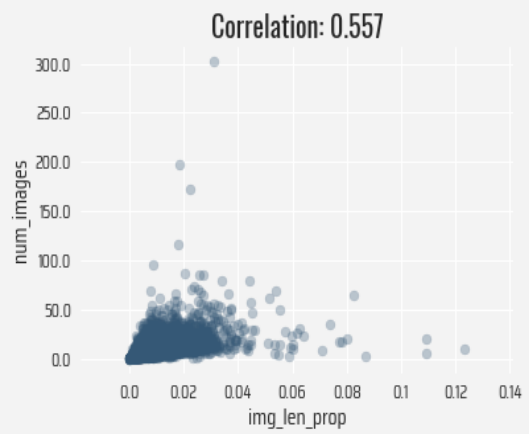- Tutorials > reporting

- But...model needs more data

THANK YOU!

APPENDIX

Pre-transform

Post-transform
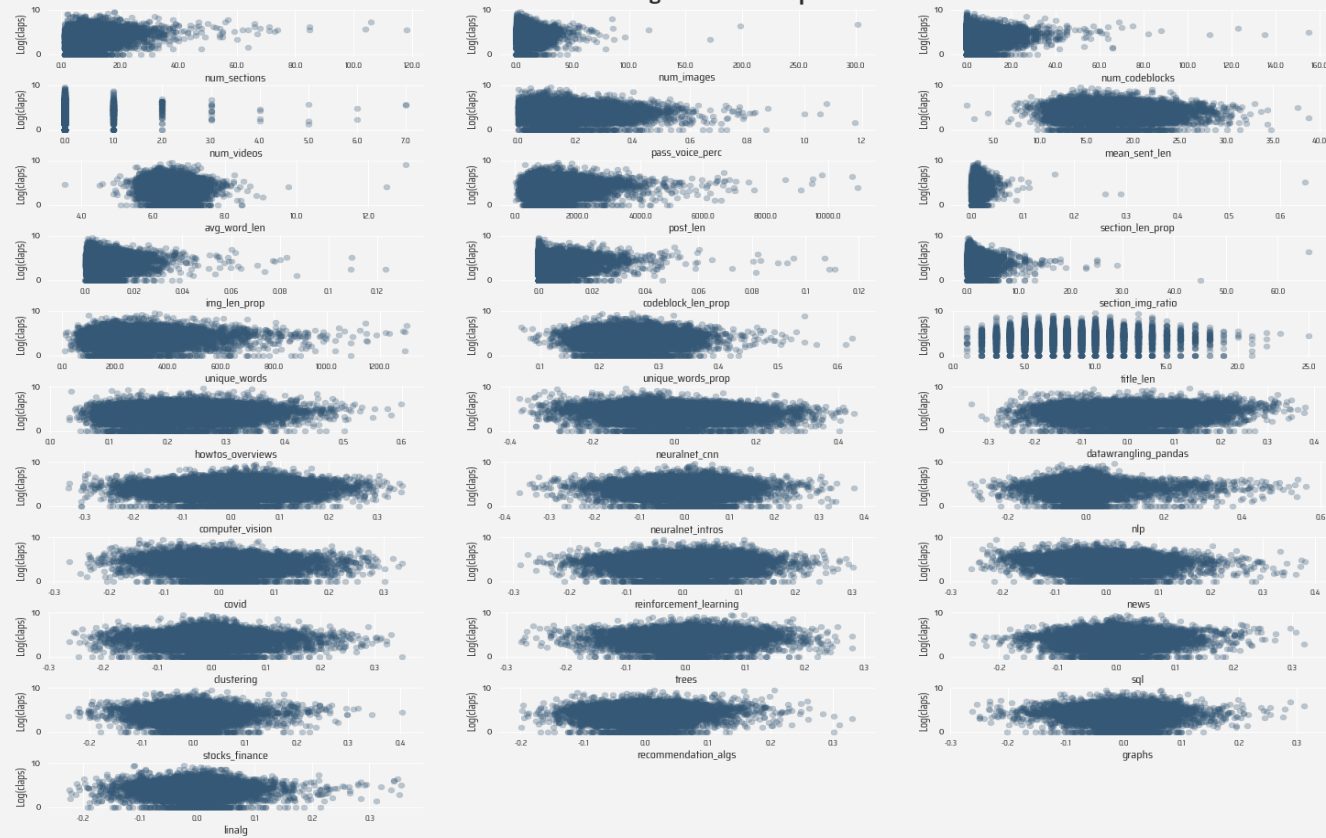
Explained Variance with Increasing Topics
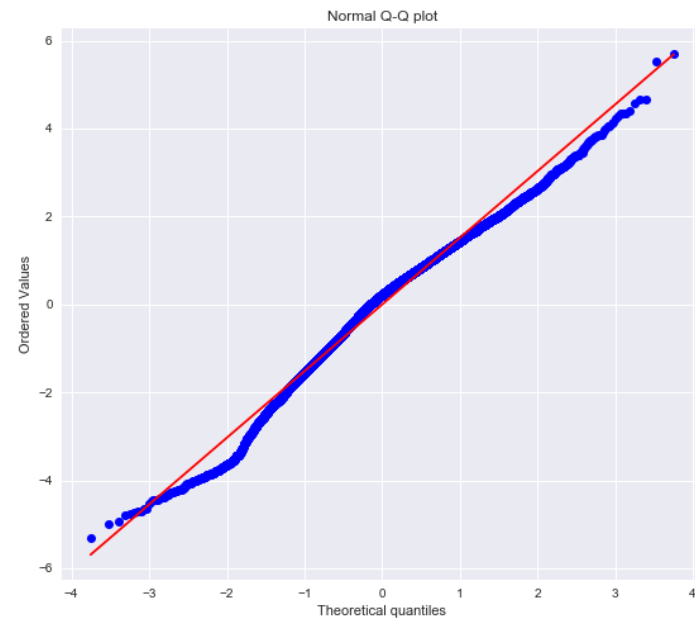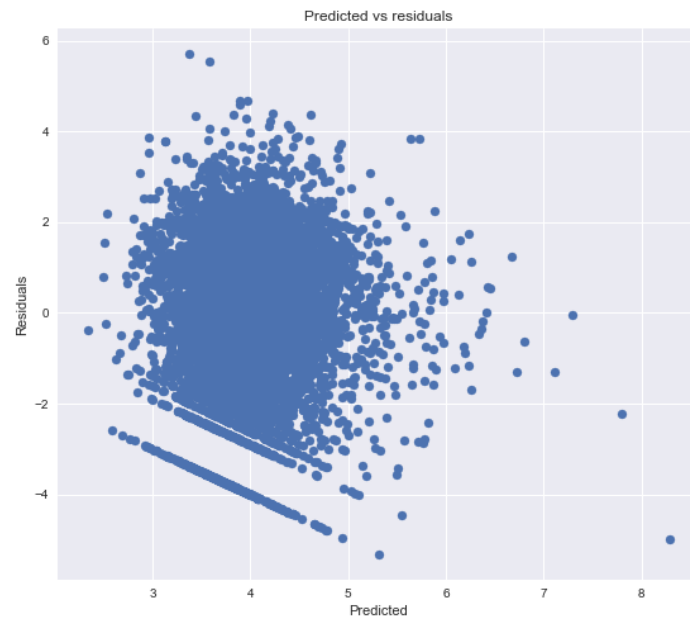
Multicollinearity Checks (r >= .5)

Feature -> Target Relationships

OLS Diagnostic Plots

CREDITS

- ◂ Presentation template by Slidesgo
- ◂ Icons by Flaticon
- ◂ Infographics by Freepik
- ◂ Images created by Freepik
- ◂ Author introduction slide photo created by Freepik
- ◂ Text & Image slide photo created by Freepik.com