

**Jose Cerecedo**  
**STA 6233**  
**Project 2**

This project is a continuation of my work from project 1. For project 1, I focused on web scrapping data from the top 3 entrepreneurship journals websites to update the Entrepreneurship Research Productivity ranking started on 2018 by Texas Christian University.

The main goal of the ranking is to collect the articles, authors, and the institution they are affiliated with to then rank the institutions productivity in terms of entrepreneurship research at peer reviewed journals.

The selection of the 3 journals on entrepreneurship is based on the Financial Times List and the journals are “Entrepreneurship Theory and Practice”, “Journal of Business Venturing: and “Strategic Entrepreneurship Journal”.

For project 2, I will use all the articles titles from the same dataset and do a text mining analysis with Shiny Word Cloud to have a clear view of what are the most common themes in each journal by analyzing each individually and in the field of entrepreneurship by analyzing all titles from the 3 journals together.

**Comparative Table of Top Words**

<b>All Titles</b>		<b>ETP Titles</b>		<b>JBV Titles</b>		<b>SEJ Titles</b>	
Venture	168	Family	117	Venture	85	Venture	34
Family	131	Firms	64	Entrepreneurs	56	Business	24
Social	101	Venture	49	New	51	Performance	24
New	100	Social	44	Social	40	Capital	23
Performance	97	Business	39	Performance	38	Strategic	22
Capital	95	Capital	37	Capital	35	New	19
Firms	93	New	36	Role	34	Social	17
Entrepreneurs	91	Performance	35	Research	23	Model	16

[illegible]

## A word cloud featuring the following terms: performance (green), new firms (purple and blue), venture (large, dark blue), family (large, dark blue, rotated), social capital (yellow and pink), and entrepreneurs (dark blue).

### ETP Titles Top 100 Words



## ETP Titles Top 8 Words





## JBV Titles Top 100 Words



## JBV Titles Top 8 Words



### SEJ Titles Top 100 Words



### SEJ Titles Top 8 Words



## **Conclusions**

From this text mining analysis we can observe some clear patterns for the entrepreneurship research literature in the top 3 peer reviewed journals and also specific patterns for individual journals.

The words “venture”, “social”, “new”, “performance” and “capital” were top words when analyzing each journal and when analyzing all journals together. It is clear these are all common trends across journals and the entrepreneurship research literature.

The words “family” and “firms” were top words when analyzing all the journals together but only top words on the Entrepreneurship Theory and Practice journal which gives us a clear indication that family business and family firms is one of the main areas of focus of this journal that differentiates it from the other two.

The word “entrepreneur” was also a top word in the analysis of all the titles from the three journals but it was only a top word in the Journal of Business Venturing, which could indicate a stronger focus of this journal on the entrepreneur as a person and at an individual level. Whereas the other two journals might have a focus more in entrepreneurship as a process and in the institutions that are part of the entrepreneurship process.

A future continuation of this project could be analyzing the abstracts of each article in order to take a deeper look into what is being researched. Another possible avenue could be doing the same analysis in different time periods and see if and how the themes evolve across time in each individual journal and in the field of entrepreneurship as a whole.

A total of 1126 articles titles were analyzed from 2015 through 2021. 218 from SEJ, 439 from JBV and 469 from ETP. The number of articles on each journal varies according to the issues and volumes they publish.

The data for this analysis can be found here: [https://github.com/ja-cerecedo/wordcloud\\_files](https://github.com/ja-cerecedo/wordcloud_files)