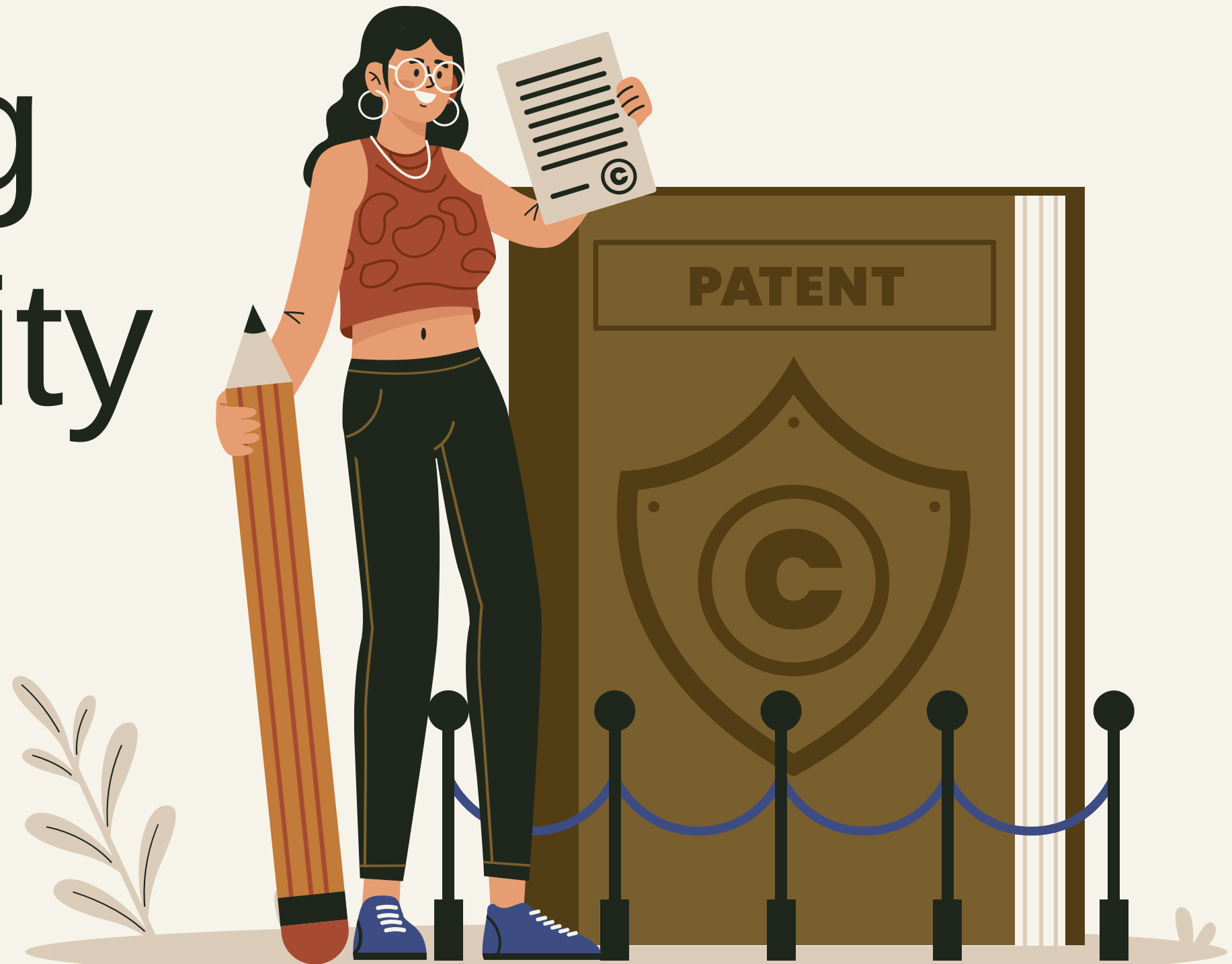


Demystifying Patent Activity

Team 01

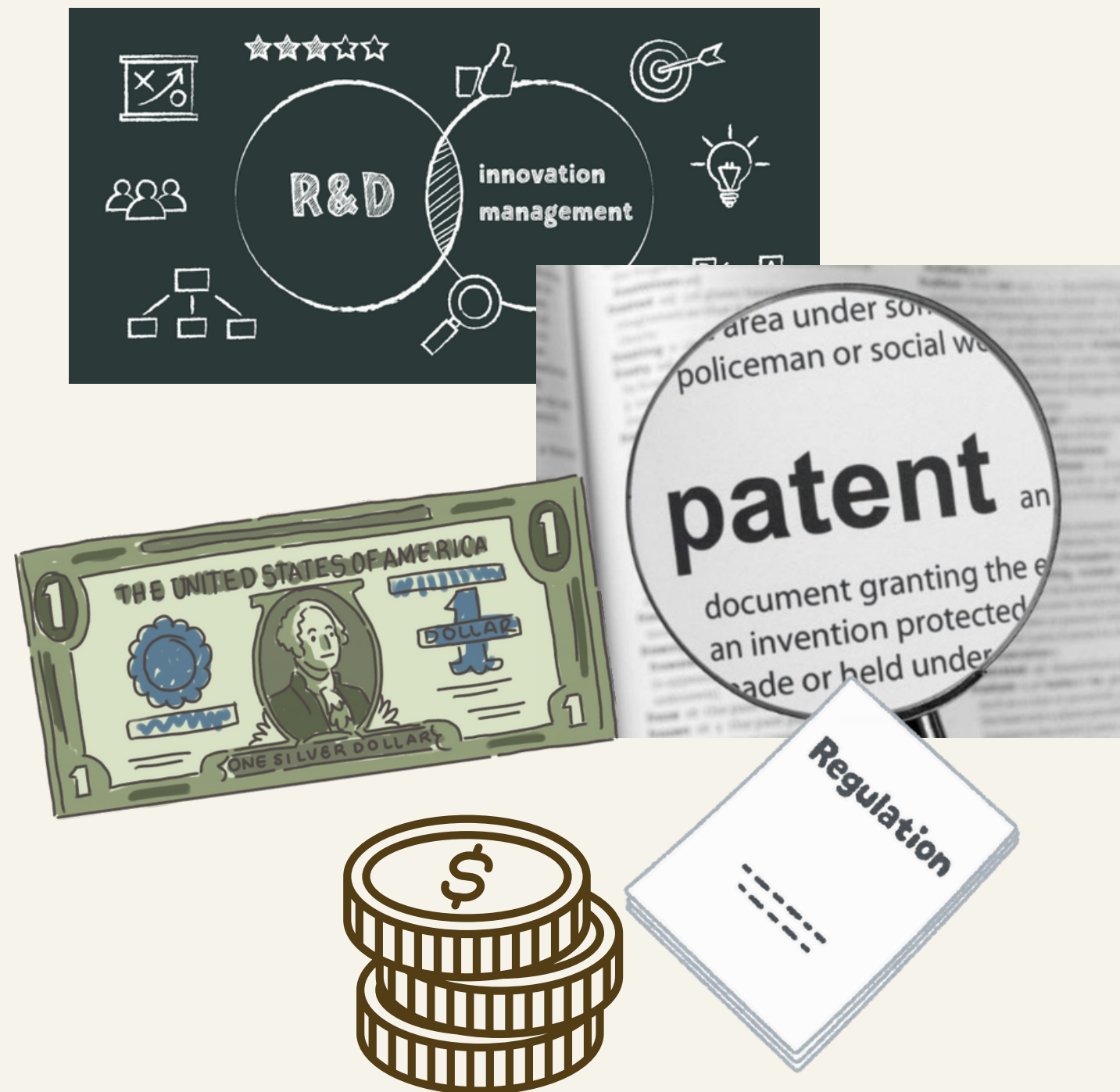
Gunjan Sharma
Jasmine Gohil
Jenil Shah
Mahika Bhartari
Sneha Ekka



Motivation



Motivation



Examining the key Internal Drivers of Company Performance in relation with patents

- Expenditure on Innovation and Research & Development
- Patent Type
- Year patent was applied for
- Industry of the company

Key stakeholder pressures can be critical drivers of a company’s performance

Competitors	Navigating Regulatory Compliance	Growth Through Consumer Adoption
Innovation Race	Pressures from the firm’s regulatory environment	Consumer response from the firm’s original and influential innovations
Benchmarking	Specific regulatory standards governing patents, product safety, environmental impact.	Innovative products drive revenue, market share, and loyalty.

Objective



Objective



Key Internal Drivers of Company Performance

- Number of Patents and activity
- Return on Assets (Revenue & Profitability) based on financial ratios

Dataset & Features



Dataset & Features

Data Source 1: Compustat Company Financials Yearly Subset

- **Features:** **gvkey**, fyear, total assets, cost of goods sold, net income, total revenue, **R&D expenses**, city, **gsector**, **sic**

Data Source 2: WRDS IBES Company Financials Yearly Subset

- **Features:** **gvkey**, year, **return on assets**, cash conversion, asset turnover, **R&D to sales**, **staff to sales**

Patent Datasets: 2 datasets that contained Patent Activity and Yearly Summary of Patent Data

Dataset & Features

Dataset 1: Patent Activity With GVKEY (subset)

PATNUM	Number of patents produced by a company in a given year
PTYPE	Type of Patent: Utility & Design
FUNCTIONAL_INNOVATIONS	Innovations by a company that were granted and protect the functional aspects of an invention.
DESIGN_INNOVATIONS	Innovations by a company that were granted and protect new, original, and ornamental designs for an article of manufacture.
GSECTOR	Industry classification

Date Range: 1990 - 2020 | Total Rows: 421182 | Columns: 27

Dataset & Features

Dataset 2: Patent Activity With GVKEY YEARLY STATS (subset)

YEAR	The year the patent was applied for
BACKWARD_CITES	An indicator of a company's level of originality in its patented innovations.
FORWARD_CITES	An indicator of a company's degree of influence in its patented innovations.
COMPANY NAME	Name of the company
COUNTRY	Name of the country

Date Range: 1990 - 2020 | **Total Rows:** 421182 | **Columns:** 27

Descriptive Stats



Industries Represented in the Dataset (gsector)

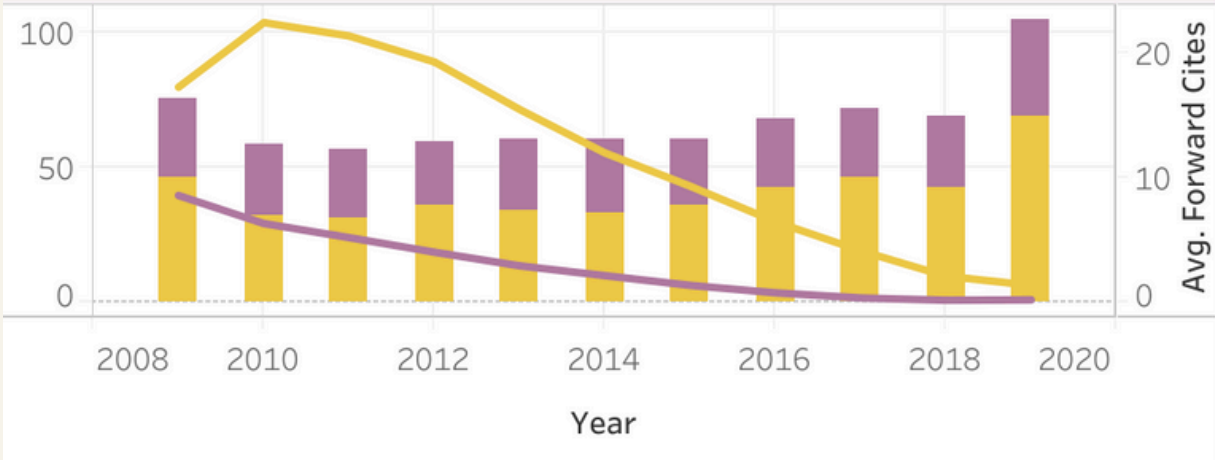
	Count	Percentage
gsector		
Information Technology	246408	58.503925
Industrials	66271	15.734528
Health Care	40188	9.541718
Consumer Discretionary	34408	8.169390
Materials	12837	3.047851
Consumer Staples	10129	2.404899
Energy	9605	2.280487
Communication	1124	0.266868
Financials	197	0.046773
Real Estate	15	0.003561

Data Viz

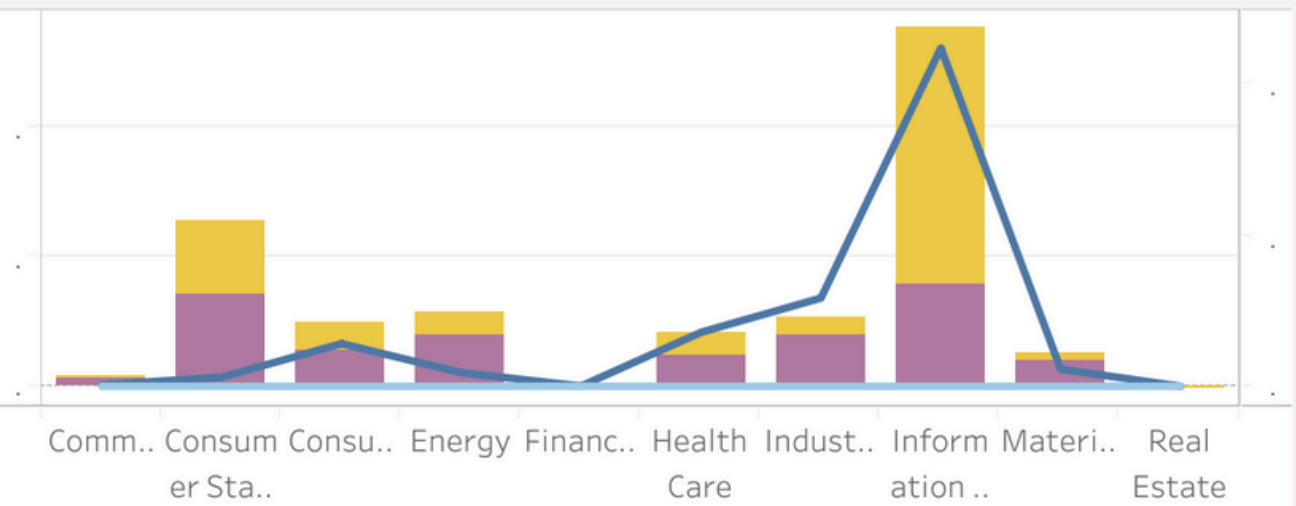


TABLEAU DASHBOARD

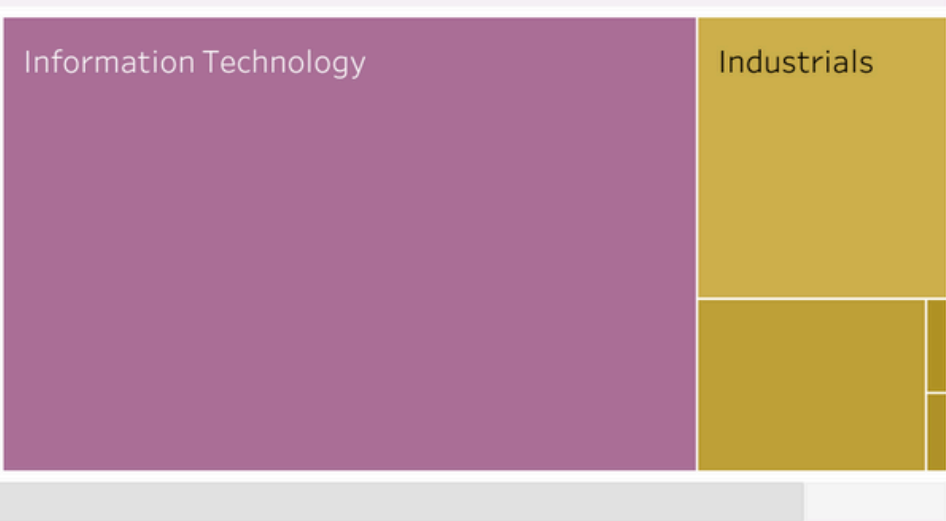
Avg No. of Backward and Forward cites across years



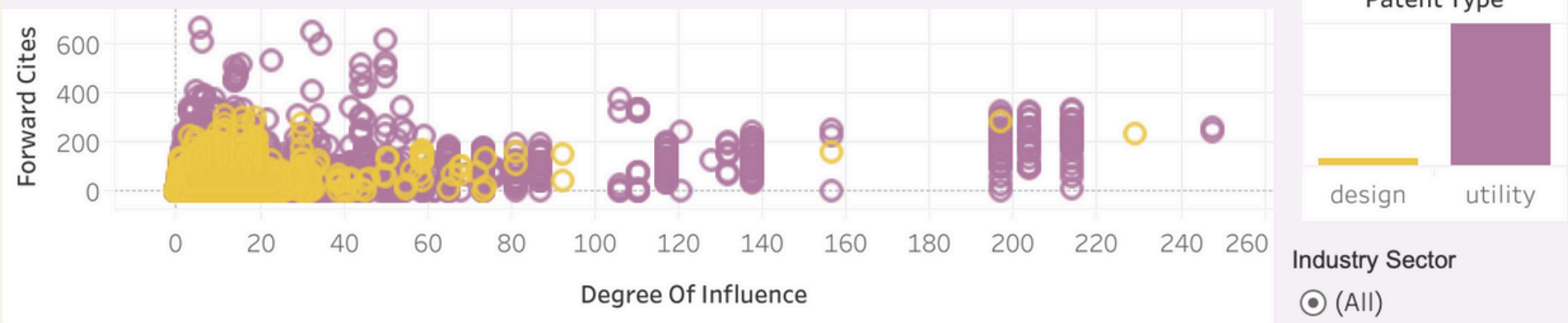
Avg Net Income and No. of patents across sectors



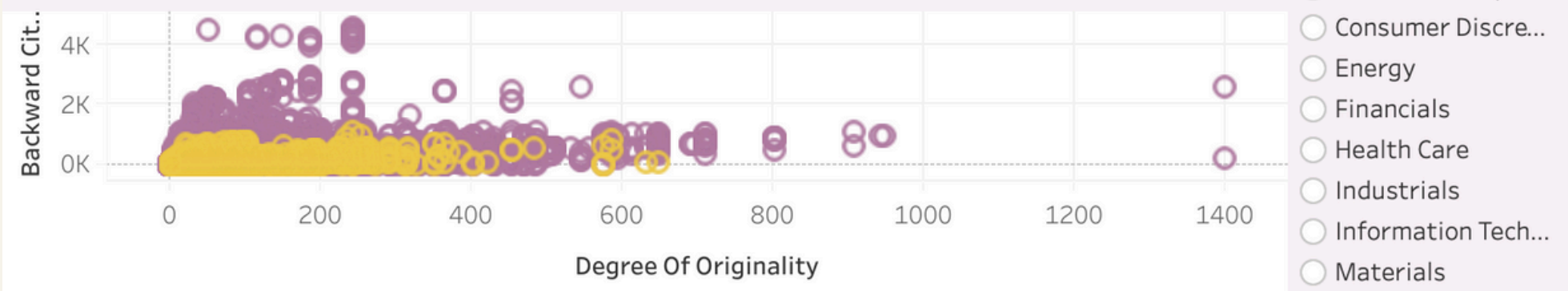
Number of Patents by Sector



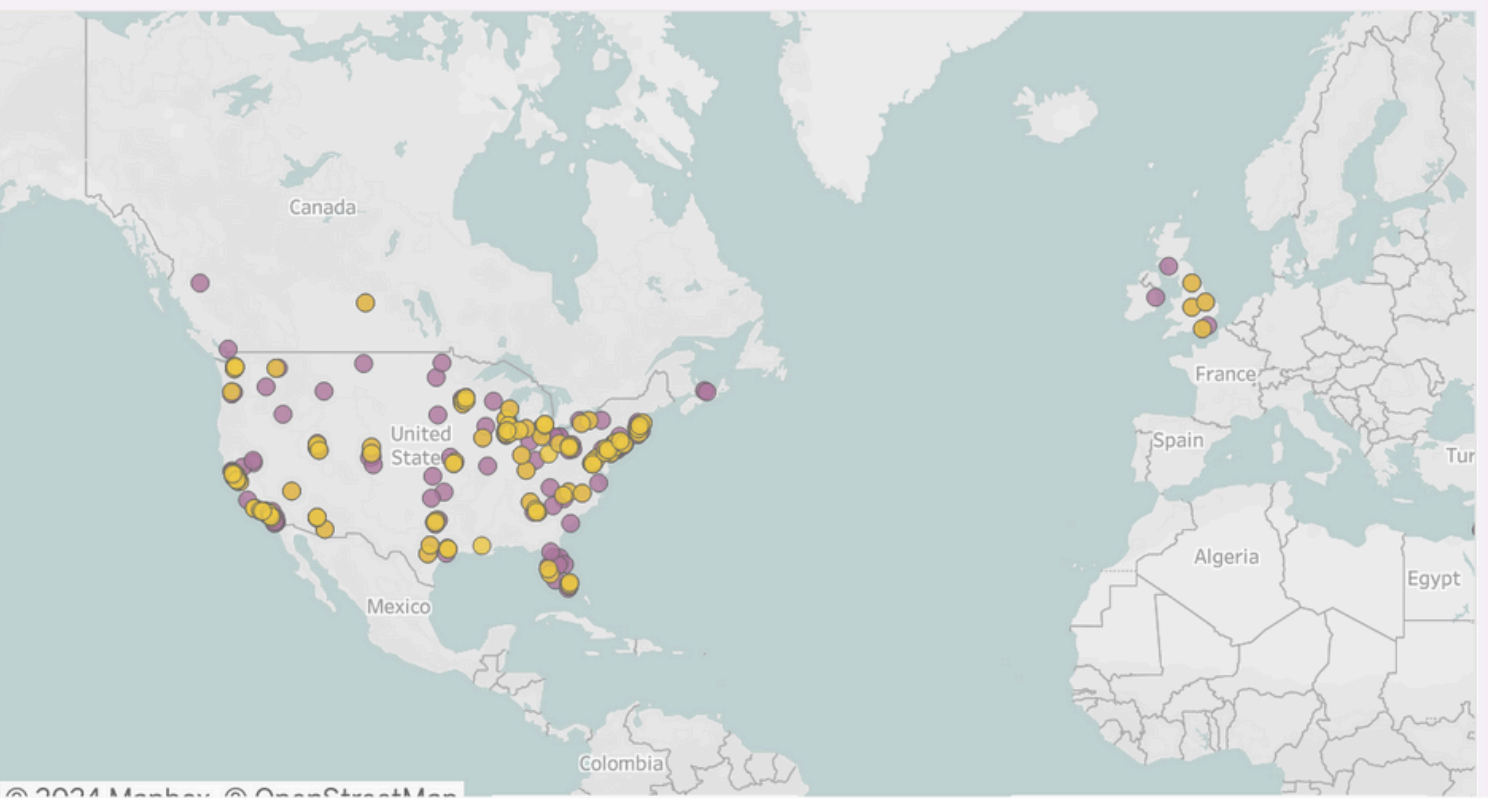
Forward Citations and Degree of Influence



Backward Citations and Degree of Originality



Number and Patent Type by cities



Regression Results



Descriptive Stats of Outcome Variables

Variable	Obs	Mean	SD	Min	Max
Return on Assets	421182	0.14	0.14	-10.37	1.63
Functional Innovation	421182	1387	2187	1.00	8750
Design Innovation	421182	1387	2187	1.00	8750

Regression (OLS) Results: DV = ROA

DV: roa	Baseline (No FE)	FE on Sector	FE on Sector and Year
Intercept	0.145*** (0.000)	-	-
backward_cites	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
forward_cites	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
degree_of_originality	-0.020*** (0.000)	0.000 (0.000)	0.000*** (0.000)
degree_of_influence	0.002*** (0.000)	0.002*** (0.000)	0.001*** (0.000)
rd_sale	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
staff_sale	-0.009*** (0.000)	-0.008*** (0.000)	-0.008*** (0.000)
year	-	-	X
sector	-	X	X
R-squared	0.017	0.080	0.086
S.E. type	iid	hetero	hetero
Observations	421182	421182	421182

Regression (OLS) Results: DV = Functional Innovations

DV: functional patents	Baseline (No FE)	FE on RD_Sale	FE on RD_Sale and Year
Intercept	1818.285 *** (7.772)	-	-
backward_cites	0.010 (0.043)	0.011 (0.025)	0.008 (0.023)
forward_cites	-0.036 (0.288)	-0.038 (0.189)	-0.039 (0.170)
degree_of_originality	-11.158 *** (0.104)	-11.157 *** (0.125)	-11.793 *** (0.130)
degree_of_influence	-22.750 *** (0.650)	-21.951 *** (1.096)	-7.579 *** (0.883)
at_turn	-71.154 *** (8.741)	-142.159 *** (4.468)	-108.728 *** (4.456)
staff_sale	-25.898 *** (3.212)	-60.859 *** (12.548)	-46.070 *** (10.593)
year	-	-	X
rd_sale	-	X	X
R-squared	0.047	0.052	0.089
S.E. type	iid	hetero	hetero
Observations	421182	421182	421182

Regression (OLS) Results: DV = Design Innovations

DV: design patents	Baseline (No FE)	FE on RD_Sale	FE on RD_Sale and Year
Intercept	1818.285 *** (7.772)	-	-
backward_cites	0.010 (0.043)	0.011 (0.025)	0.008 (0.023)
forward_cites	-0.036 (0.288)	-0.038 (0.189)	-0.039 (0.170)
degree_of_originality	-11.158 *** (0.104)	-11.157 *** (0.125)	-11.793 *** (0.130)
degree_of_influence	-22.750 *** (0.650)	-21.951 *** (1.096)	-7.579 *** (0.883)
at_turn	-71.154 *** (8.741)	-142.159 *** (4.468)	-108.728 *** (4.456)
staff_sale	-25.898 *** (3.212)	-60.859 *** (12.548)	-46.070 *** (10.593)
year	-	-	X
rd_sale	-	X	X
R-squared	0.047	0.052	0.089
S.E. type	iid	hetero	hetero
Observations	421182	421182	421182

Summary



SUMMARY

All independent variables exhibit statistical significance with p-values below 0.05, indicating that they have an influence on the dependent variables. However, the observed impact on the dependent variables is relatively small.

Takeaways



KEY TAKEAWAYS

- The Variance Inflation Factor (VIF) was found to be below the threshold of 10 when Return on Assets (ROA) was considered as the dependent variable.
- The Variance Inflation Factor (VIF) exceeded acceptable levels in the models incorporating functional and design innovations as dependent variables. Consequently, these variables were not utilized as dependent variables in the analysis due to multicollinearity concerns.

Challenges



CHALLENGES

- Numerous null values contained in Financial Ratios that could have helped explain the variance in the dependent variables.
- Computational Costs were high
- Design & functional innovations = high collinearity

Thank You

