Venture Funding & Patent Portfolio

PitchBook

Then

Dan, VC Investor

Dan goes to PitchBook to find companies to invest in

PitchBook gives Dan two promising companies

How should Dan select a company if their financials are seemingly equivalent?

Dan would look at each company's patent portfolio which is available on US Patent Database (USPTO)

Problems with Dan's Approach

1. Large USPTO Dataset: **81M Rows Of Data**

2. Computationally Intensive: **32 Billion Possible Queries**

3. Textual Data : **No Visualizations**

According to a Harvard Business Research Study, Venture Capital Investors spend about **20%** of their time trying to find relevant company data

Our Project

Extract USPTO Data

Use n-grams, TF-IDF Vectorizing & Nearest Neighbors Algorithm to match data with PitchBook's company data
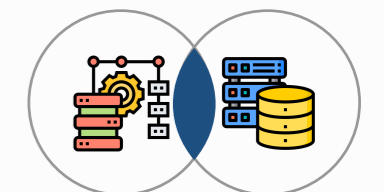
Use embedded SQL querying to dynamicslly pull matched companies' patent information

Merge resulting data to be used for gaining insights into companies' patent portfolio on PitchBook
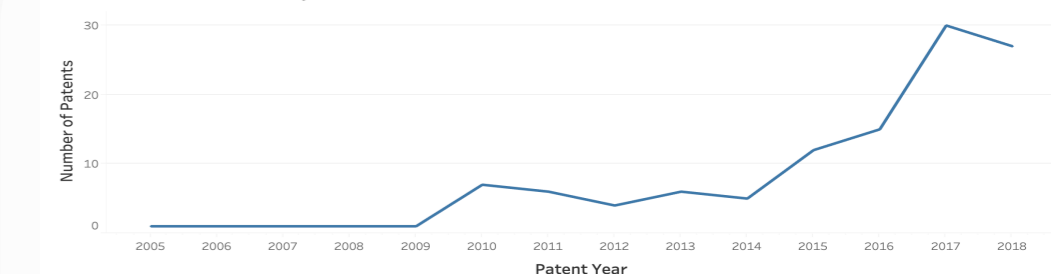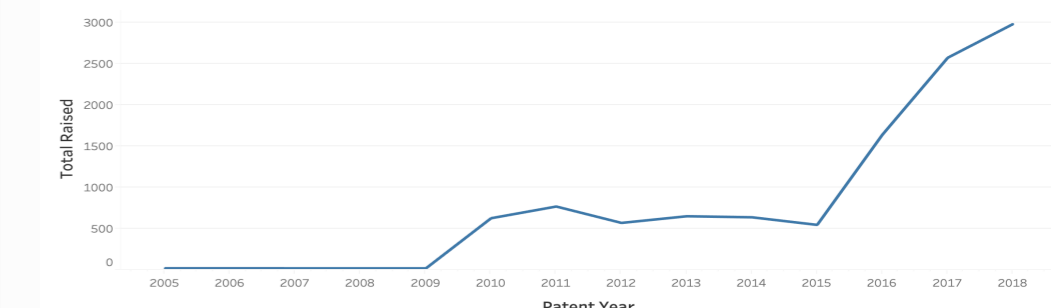
Now

Dan can go to pitchbook to find companies

PitchBook will use our matching algorithm to obtain data about companies and their patent portfolio

Number of Patents Yearly

Total Raised vs Patent Year

Now Dan can look at each company's patent portfolio and trends to decide which company to invest with

Dan finally finds a company to invest in!

Arjun Goswami | Netra Pathak | Yash Patel | Rishabh Shah

Information School
UNIVERSITY *of* WASHINGTON