ANALYSIS OF CONTRACTOR DATA IN FEDERAL ACQUISITION DATABASES

Ningning Wu, M. Eduard Tudoreanu, Sanketh Siruvolu, Richard Wang, John R. Talburt

Institute for Chief Data Officers
University of Arkansas at Little Rock
Massachusetts Institute of Technology

Research Question

- A wealth of information is collected and available
  - Governmental sources
  - Social media
  - Websites
- Valuable information could be used for supporting effective acquisition planning and management
- Could more knowledge be extracted by combining acquisition data with other open data?
  - Federal acquisition and purchasing databases have rich information about past contracts

Importance to Acquisition

- Social, economic, natural changes often get captured in various data sources
  - Can provide a more global and integrated view of factors affecting contracts and projects
  - Complementary for
    - subject-matter expert knowledge
    - contractor information and expertise

Agenda – Two Topics

- Natural Disaster Risk Map
  - Difficult to assess the natural disaster risk level of an area, as there is no comprehensive classification system
  - We propose a metric that utilizes both National Centers for Environmental Information (NCEI) and Federal Emergency Management Agency (FEMA) to classify the risk levels of all U.S. counties
- Natural Disaster Risk Analysis of Past Projects
  - We determine the distribution of federal contractors and business in different risk areas

Natural Disaster Data

- We analyze natural disaster data for **each US county** between the years 1950 and 2018 from the National Centers for Environmental Information

  - Formerly the National Climatic Data Center NCDC

- All types of natural disasters, including flood, tornado, hurricane, blizzard, high wind, flash flood, hail, dust storm, etc.

- We produced a natural disaster map of the U.S.

FEMA Data

- We have retrieved the emergency assistance data for **each US county** between the years 1953 and 2020

- The data contains the following fields:

1	femaDeclarationString	13	incidentBeginDate
2	disasterNumber	14	incidentEndDate
3	state	15	disasterCloseoutDate
4	declarationType	16	fipsStateCode
5	declarationDate	17	fipsCountyCode
6	fyDeclared	18	placeCode
7	incidentType	19	designatedArea
8	declarationTitle	20	declarationRequestNumber
9	ihProgramDeclared	21	hash
10	iaProgramDeclared	22	lastRefresh
11	paProgramDeclared	23	id
12	hmProgramDeclared		

Weighted Disaster Score(WDS)

- WDS is a weighted sum of assistance programs received by an area between 1953 and 2020

- It aims to infer an area's risk level and is defined as follows:

$$s = \sum_{i=1}^{4} w_i \times n_i$$

Where n_i is the number of a specific type of the FEMA assistance programs, and w_i is the corresponding weight for the type

- The weight for each assistance program of WDS score is defined as follows:
  - Disaster mitigation: 0.25
  - Public assistance : 0.50
  - Housing assistance: 0.75
  - Individual assistance: 1.0

Five-Number Summary of WDS and Number of Disasters for All U.S. Counties

Statistic	WDS	# of Disasters
Min	0.25	1
Q1	6.25	10
Medium	10	15
Q3	13	19
Max	56.25	105

◉ The first quartile and the third quartile of WDS are used as the cutoff points for the risk classes as shown below:

level	WDS values
Low	wds<1st quartile
Medium	1st quartile <wds<3rd quartile
High	wds> 3rd quartile

Geographical Distribution by Risk Levels

Risk Distribution of US Counties
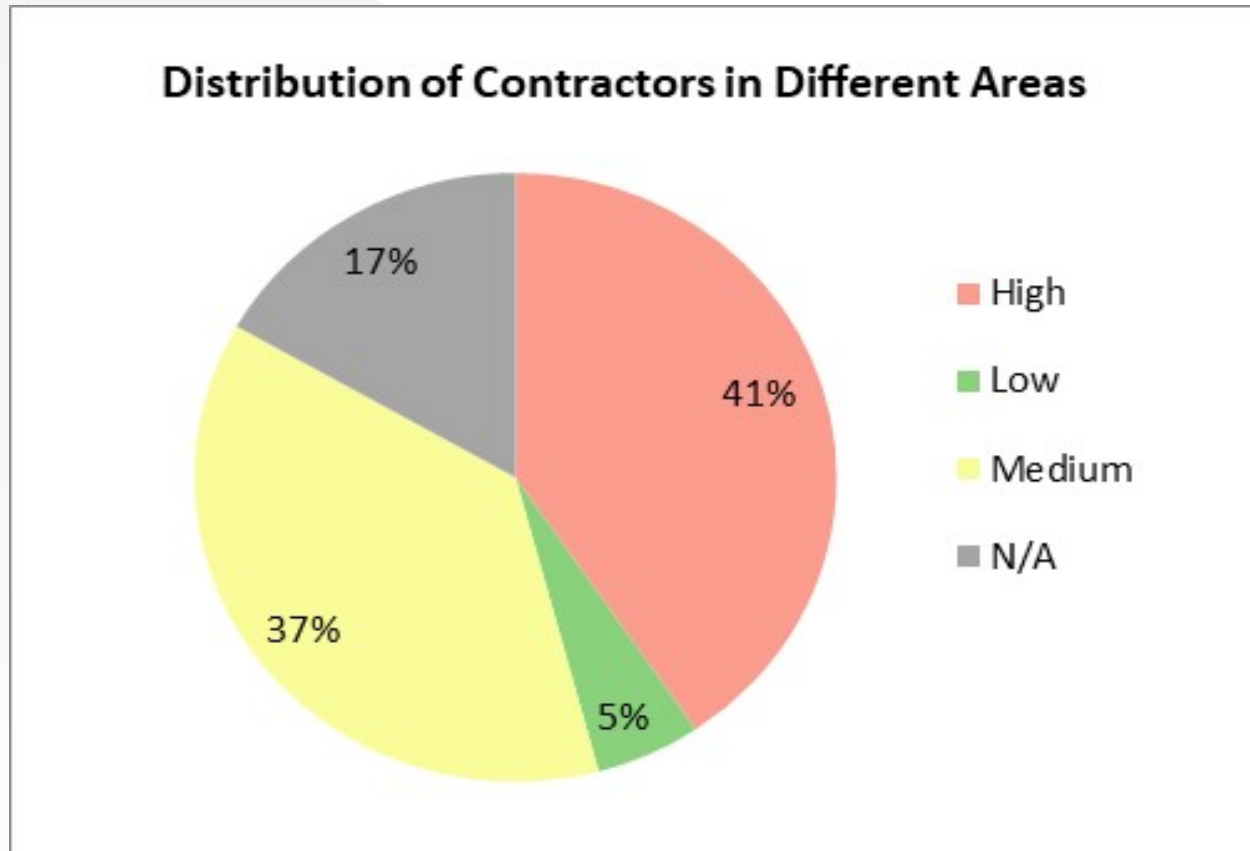
- low: 20%
- medium: 52%
- high: 28%
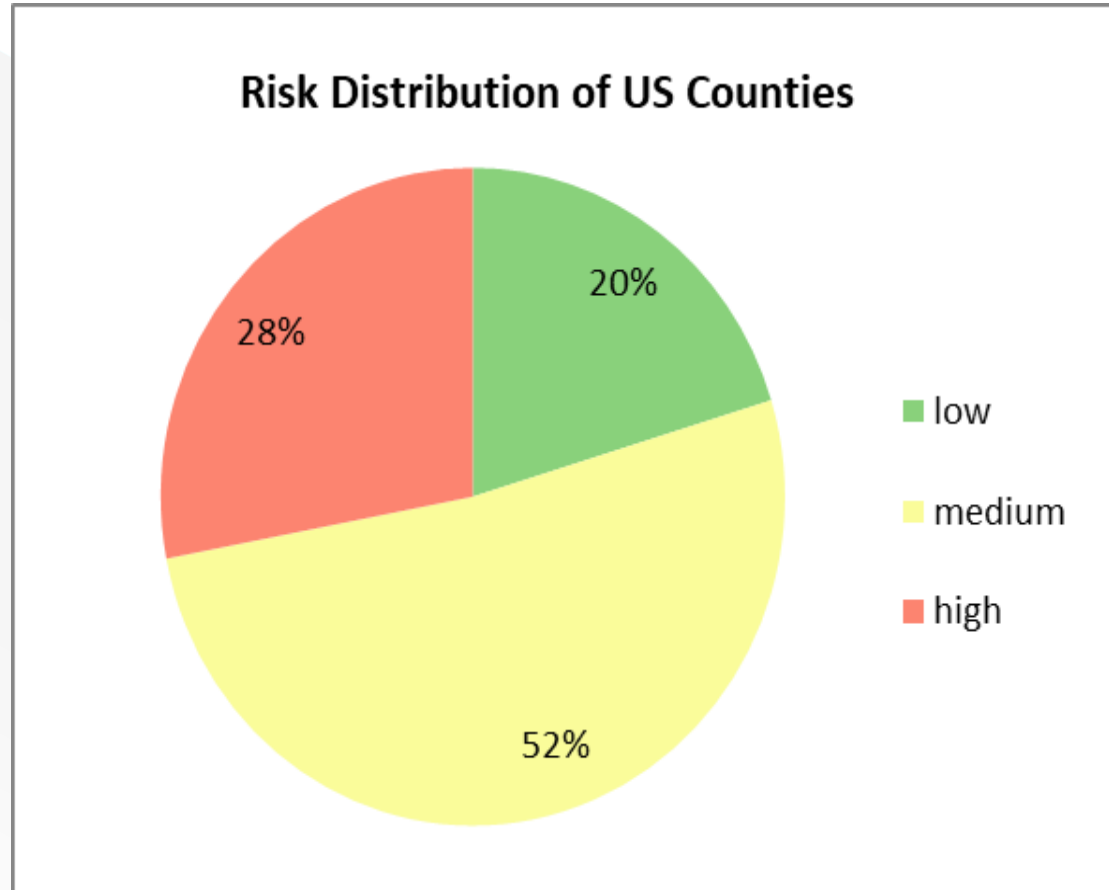
Natural Disaster Risk Map for U.S. Counties

- The following figure shows Natural disaster risk class (low, medium, high) displayed for each county in the United States. Red encodes high risk, yellow medium, and green low.
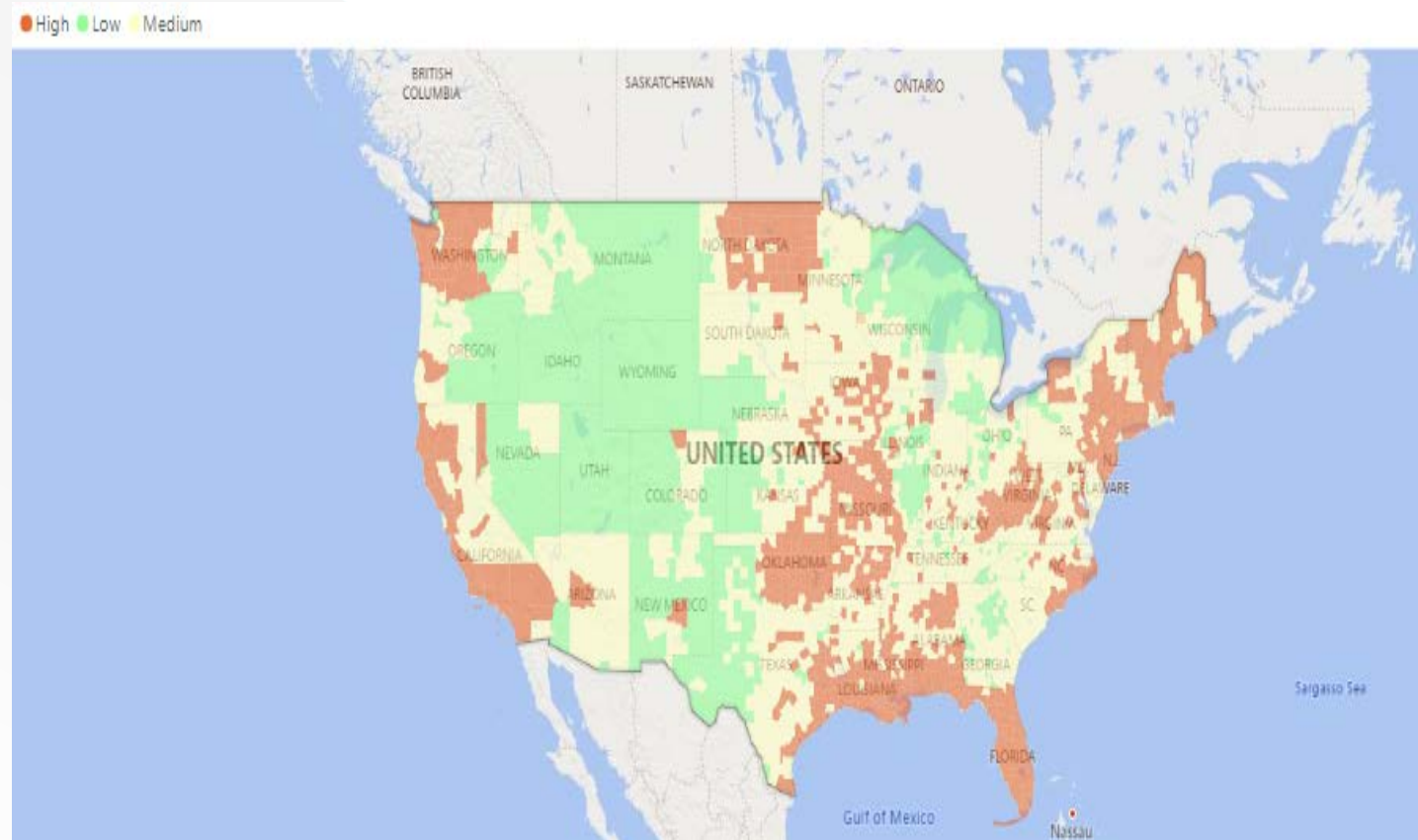
# Federal Contractors Distribution by Risk Levels

The figure shows the distribution of federal contractors divided by class of natural disaster risk

- As high as 41% of contractors are located in high-risk areas.

- About 17% of contractors are not located in U.S. so their natural disaster risk levels are not assessed by this research.
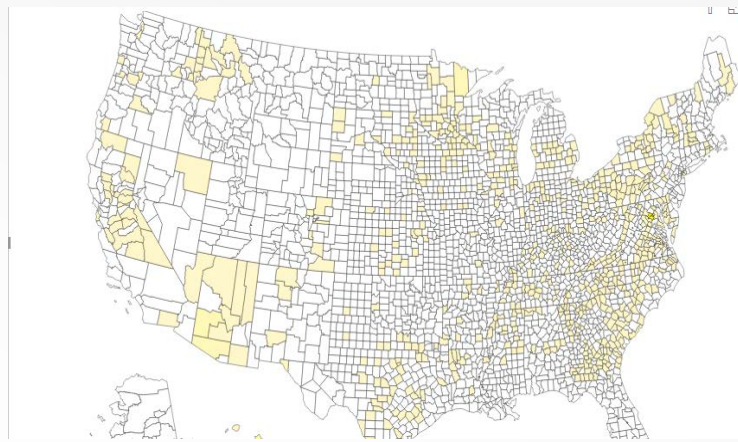
Place of Performance of Navy Awards Correlated to Natural Disaster Risk
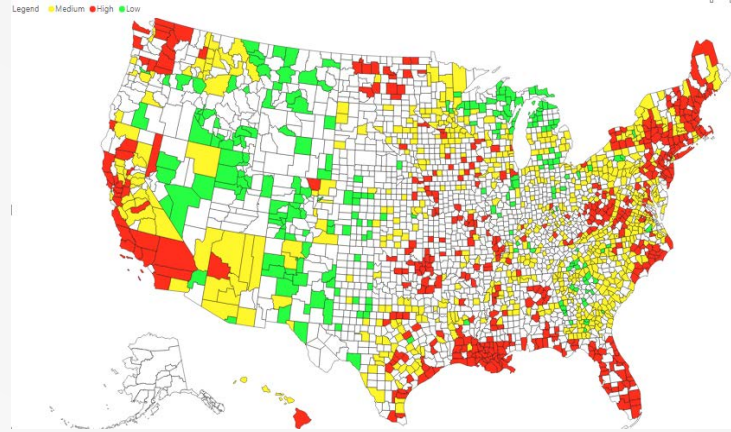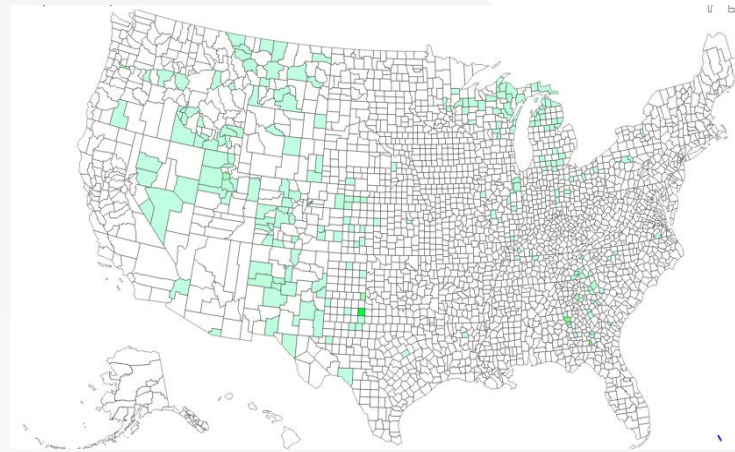

(a)
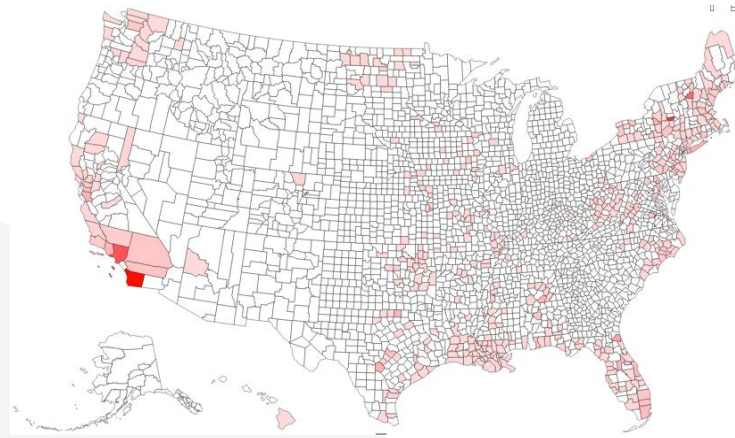
(b)

(c)

(d)

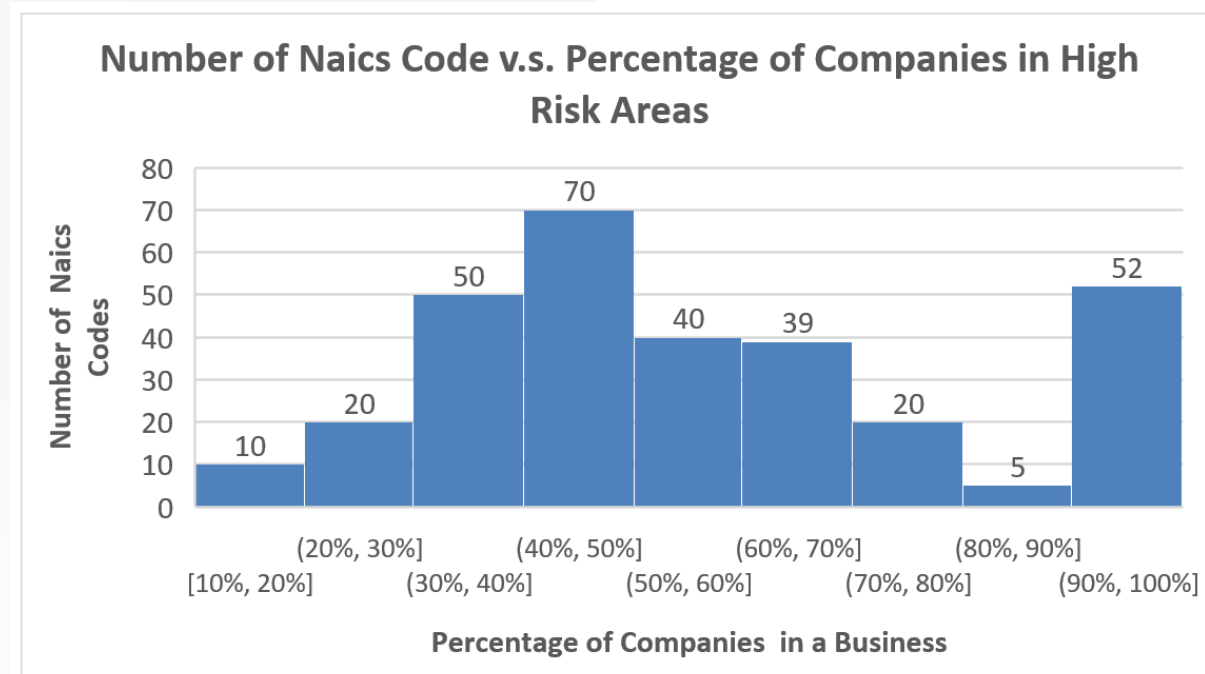(a) shows overall view of counties with at least one award.

(b) shows only counties with high risk using color intensity to encode the number of awards in that county

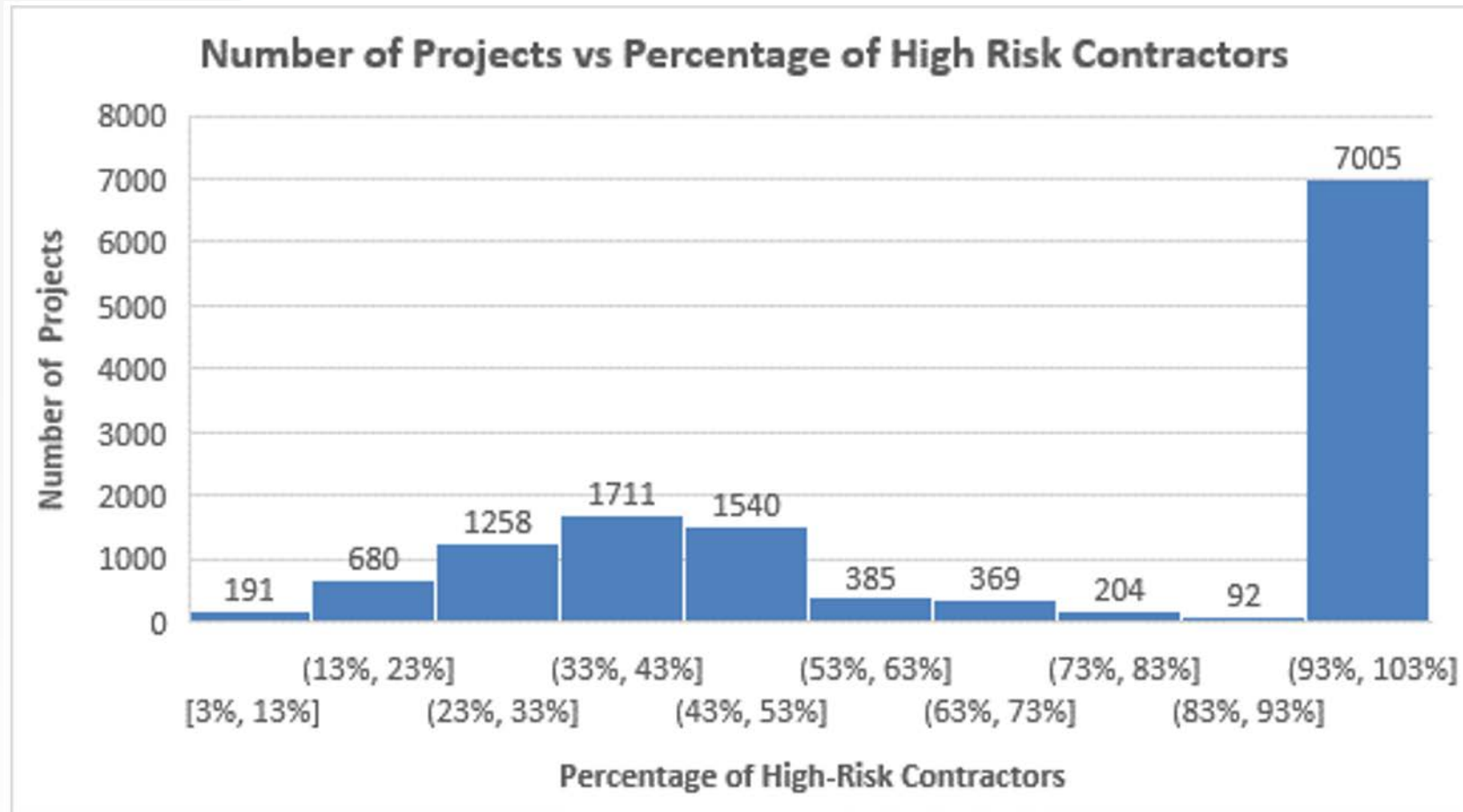(c) depicts counties with medium risk level, using color intensity for number of awards

(d) shows low risk counties and number of awards.

# Clusters of NAICS Codes by the Percentages of High-Risk Companies

- There are 353 unique NAICS codes of all contractors
- The clusters of NAICS codes based on the percentage of companies located in high-risk areas is shown as below:
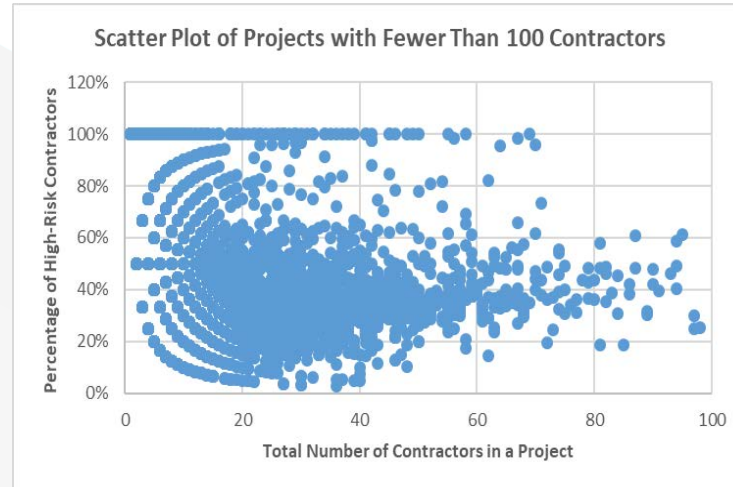
Number of Naics Code v.s. Percentage of Companies in High Risk Areas

Number of Naics Codes

Bin	Count
[10%, 20%]	10
(20%, 30%]	20
(30%, 40%]	50
(40%, 50%]	70
(50%, 60%]	40
(60%, 70%]	39
(70%, 80%]	20
(80%, 90%]	5
(90%, 100%]	52

Percentage of Companies in a Business

Percentage of High Risk Contractors for All Past Projects

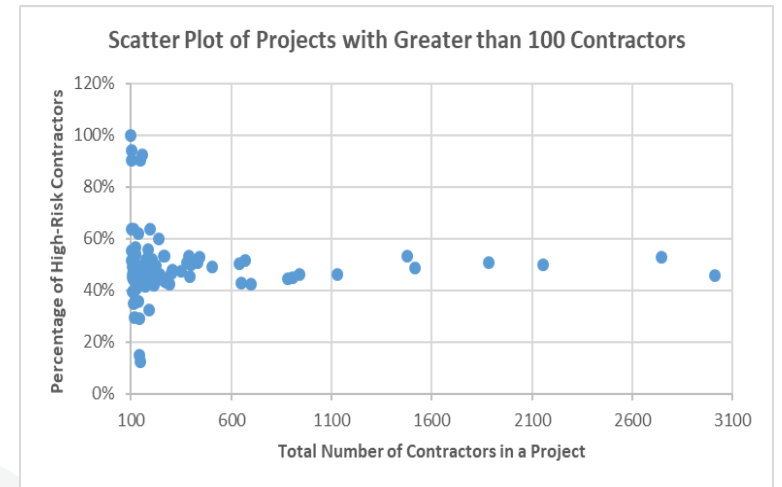# Total Number of Contractors in a Project vs. Percentage of High Risk Contractors
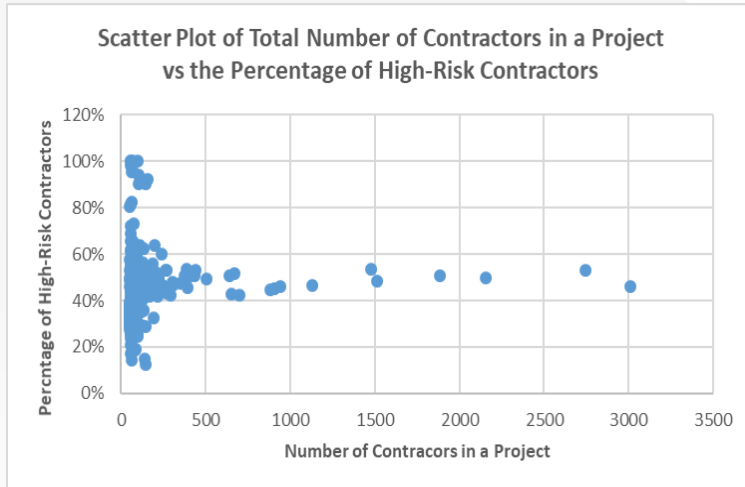
(a)   (b)   (c)

(a) shows the scatter plot of the total number of contractors in a project and the percentage of the high-risk ones. It shows that the majority of projects have fewer than 250 contractors. There are several projects with more than 1000 contractors.

(b) and (c) shows the scatter plot for the projects with more than 100 contractors and the ones with fewer than 100 contractors respectively.

Concluding Thoughts

- Unique service contractors might increase the risk associated to an acquisition project, especially if they are located in areas with high occurence of natural disasters

  - Need strategic plan to manage project if/when the contractors become unavailable

- Information on high risk areas of natural disasters is beneficial as it would help project managers assess the risk to project timeline and develop strategies to mitigate the risk