

An Interdisciplinary Approach to Data Science

Galin Jones

University of Minnesota

July 2018

Data Science at Minnesota

In 2012 faculty from

Statistics

Biostatistics

Computer Science

Electrical Engineering

Industrial Engineering

Health Informatics

Astrophysics

Mathematics

...

met to start thinking about a Master's degree program in data science.

What Is Data Science?

What Is Data Science?

We didn't know

What Is Data Science?

We didn't know

and there was no consensus view.

What Is Data Science?

We didn't know

and there was no consensus view.

There wasn't even a majority view.

What Is Data Science?

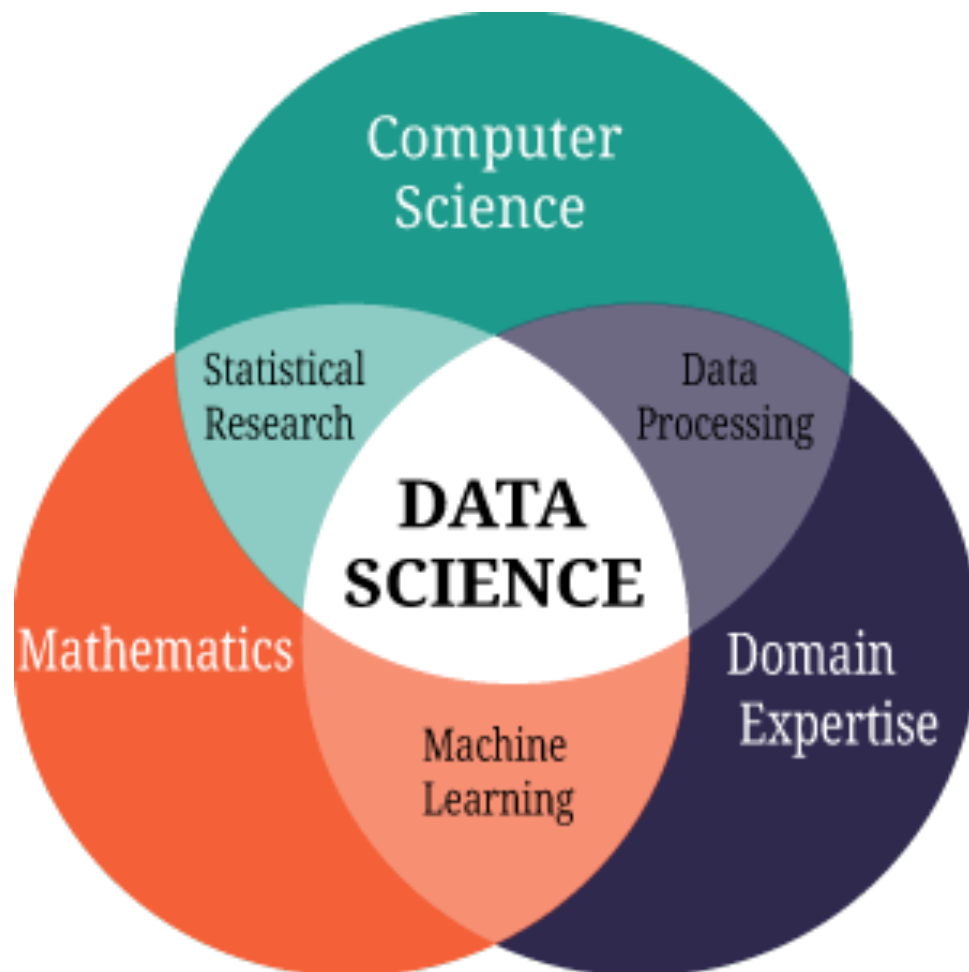


Figure 1: quora

What Is Data Science?

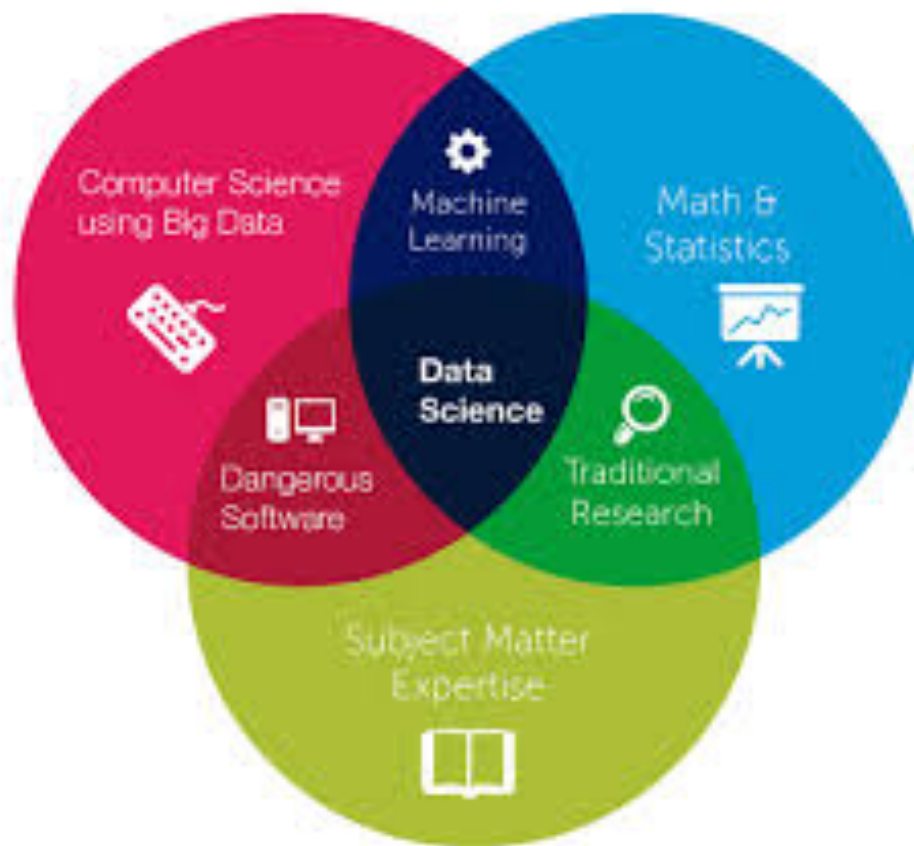


Figure 2: techiexpert.com

What Is Data Science?



Figure 3: towarddatascience.com

What Is Data Science?

We still don't know, but we decided to take an inclusive view.

What Is Data Science?

We still don't know, but we decided to take an inclusive view.

What Is Data Science?

We still don't know, but we decided to take an inclusive view.

Consultation with MinneAnalytics was a key step.

In addition to traditional statistics, some data scientists routinely engage in

- Data collection
- Data archiving
- Software development
- Visualization
- ...

What Is Data Science?

We still don't know, but we decided to take an inclusive view.

Consultation with MinneAnalytics was a key step.

In addition to traditional statistics, some data scientists routinely engage in

- Data collection
- Data archiving
- Software development
- Visualization
- ...

This led to an MS degree that has only a loose structure.

Data Science at Minnesota

Basic structure of the degree

- Statistics (2 courses)

- Algorithms (2 courses)

- Infrastructure (2 courses)

- Electives (3 courses)

- Capstone (1 course)

Students choose from a list of approved courses in each category.

Emphasis on presentation skills and internships

Prerequisites for Data Science

Prereqs	DS
Multi Calc	Yes
Lin Alg	Yes
Adv Calc	No
Gen Stat	Yes
Adv Stat	No
Prog	Yes (2)
Matlab/R	Pref
Adv Data Str	Pref

Application Pool Characteristics

DataSci	FY16	FY17	FY18
Apps	241.0	269.0	462.0
PerFem	29.5	21.9	30.5
PerIntl	92.5	94.1	93.1

Application vs Matriculation Pool Characteristics

DataSci	FY16	FY17	FY18
Apps	241.0	269.0	462.0
PerFem	29.5	21.9	30.5
PerIntl	92.5	94.1	93.1

DataSci	FY16	FY17	FY18
Mats	20	18.0	27.0
PerFem	30	11.1	40.7
PerIntl	65	61.1	92.6

Consequences

Size of MS courses has increased.

Consequences

Size of MS courses has increased.

Courses serve multiple audiences. (DS, Stat MS, Stat BS/BA)

Consequences

Size of MS courses has increased.

Courses serve multiple audiences. (DS, Stat MS, Stat BS/BA)

Course prerequisites are difficult to enforce

Consequences

Size of MS courses has increased.

Courses serve multiple audiences. (DS, Stat MS, Stat BS/BA)

Course prerequisites are difficult to enforce

Modification of existing courses

Consequences

Size of MS courses has increased.

Courses serve multiple audiences. (DS, Stat MS, Stat BS/BA)

Course prerequisites are difficult to enforce

Modification of existing courses

New course development (stat computing and stat learning)

Prerequisites: Data Science MS vs Stat MS

Prereqs	DS	STAT
Multi Calc	Yes	Yes
Lin Alg	Yes	Yes
Adv Calc	No	Pref
Gen Stat	Yes	Yes
Adv Stat	No	Yes
Prog	Yes (2)	Yes (1)
Matlab/R	Pref	No
Adv Data Str	Pref	No

Application Pool Characteristics (DataSci MS vs Stat MS)

APPS	FY16	FY17	FY18
DataSci	241	269	462
Stat	312	257	235

Application Pool Characteristics (DataSci MS vs Stat MS)

APPS	FY16	FY17	FY18
DataSci	241	269	462
Stat	312	257	235

PerFem	FY16	FY17	FY18
DataSci	29.5	21.9	30.5
Stat	49.7	49.8	49.6

Application Pool Characteristics (DataSci MS vs Stat MS)

APPS	FY16	FY17	FY18
DataSci	241	269	462
Stat	312	257	235

PerFem	FY16	FY17	FY18
DataSci	29.5	21.9	30.5
Stat	49.7	49.8	49.6

PerIntl	FY16	FY17	FY18
DataSci	92.5	94.1	93.1
Stat	84.3	88.7	90.2

What's working and what isn't

What's working

What's working and what isn't

What's working

We have a Data Science MS program.

What's working and what isn't

What's working

- We have a Data Science MS program.
- Attracting many applicants.

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

- Thinking about a BS degree.

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

- Thinking about a BS degree.

and what isn't

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

- Thinking about a BS degree.

and what isn't

- Capstone is an obstacle to scale.

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

- Thinking about a BS degree.

and what isn't

- Capstone is an obstacle to scale.

- Stat courses are more challenging to teach.

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

- Thinking about a BS degree.

and what isn't

- Capstone is an obstacle to scale.

- Stat courses are more challenging to teach.

- Unclear what the Data Science degree means.

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

- Thinking about a BS degree.

and what isn't

- Capstone is an obstacle to scale.

- Stat courses are more challenging to teach.

- Unclear what the Data Science degree means.

- Difficult to update/modify curriculum.

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

- Thinking about a BS degree.

and what isn't

- Capstone is an obstacle to scale.

- Stat courses are more challenging to teach.

- Unclear what the Data Science degree means.

- Difficult to update/modify curriculum.

- Applications to stat MS have decreased.

What's working and what isn't

What's working

- We have a Data Science MS program.

- Attracting many applicants.

- Internships

- Capstone

- Thinking about a BS degree.

and what isn't

- Capstone is an obstacle to scale.

- Stat courses are more challenging to teach.

- Unclear what the Data Science degree means.

- Difficult to update/modify curriculum.

- Applications to stat MS have decreased.

- Only one of many UM organizations engaged in Data Science.