Engaging Students in the Time of Business Analytics

Bob Stine
Department of Statistics
The Wharton School
University of Pennsylvania

Wharton
Department of Statistics

Perspective

Currently…

Teaching required undergrad regression

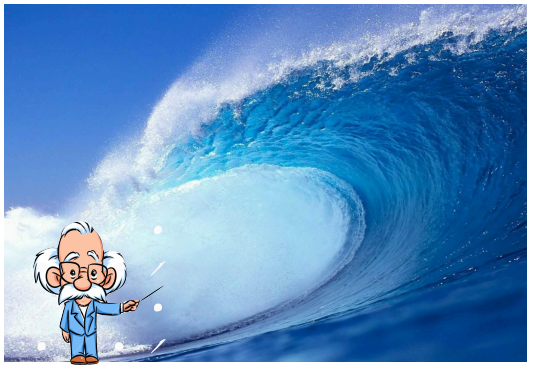
- "Analytical" regression
- Focus on answering business questions
- Emphasize motivation/interpretation/action rather than "theory"
- Software such as JMP, Excel, StatCrunch

3 sections, each with 120 students

- 1 TA

Gateway or barrier to later courses…

Business Analytics



Everywhere

- Particularly in business schools
- Elsewhere "Data Science"

At Wharton

- New joint Business Analytics concentrations
  - Undergraduate
  - MBA
- Attractive STEM field
  - Facilitates VISA
- Soaring enrollment for many courses
  - Modern data analytics
  - R programming
  - Probability… one section @ 30 has grown to 3 @ 65

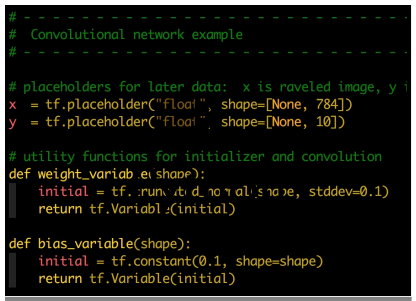
Challenges to Engagement

- Scale
  - Attention span
  - Phones
  - Behavioral changes
- Freshmen, sophomores
  - Don't know business: e.g. CAPM, elasticity, …
  - Business Analytics has expanded variation
  - Growing subset of AP students adds yet more
- Placement: Statisticians' view of Business Analytics is not the same as recruiters' view

Bus Analytics ≠ Applied Stat

Talk to those hiring Business Analytics grads

- Data manipulation
- Programming, scripting
- Problem solving, willing to dive in
- And, yes, statistical sense for data analysis



Wharton concentration

- Joint concentration with "Decision Science" OID
- Deliberate distinction from Stat concentration
  - Probability course is not in the concentration!
- Mixture: Data collection, Data analysis, Optimization

Lesson: Collaboration enhances marketable skills

Lack Business Knowledge

Undergraduate business program

- Revised undergrad curriculum
- First year exploration
- Second and third year skills
- Senior year capstone

Capstone course

- Limited size, group project
- Blend knowledge from various courses
- Application field project

Hard to find, harder to sustain over time…

Lesson: More collaboration!

Engagement at Scale

- Compelling examples are key
  - Distract from those phones as well…
  - Compelling for MBA ≠ Compelling for Undergrad
- Topical examples
  - Themselves
  - Contemporary issues
- Software
  - Not teaching programming
    - At least not at the introductory level!
  - Jazzy animation

Examples that click

- Student generated data
  - Number text messages, male vs female
  - Hard to manage (much less grade) at scale

projects

- Substantive examples
  - Contemporary: Election surveys, sports
  - Climate change
  - Crime rates
  - Wage discrimination
- "Eye candy" examples
  - Macroeconomic data

Election Surveys

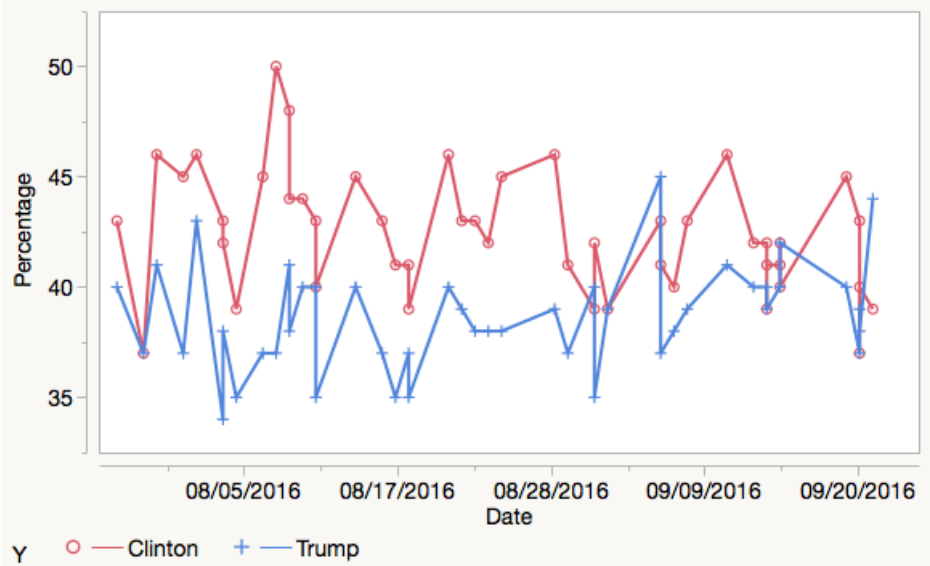
Topic

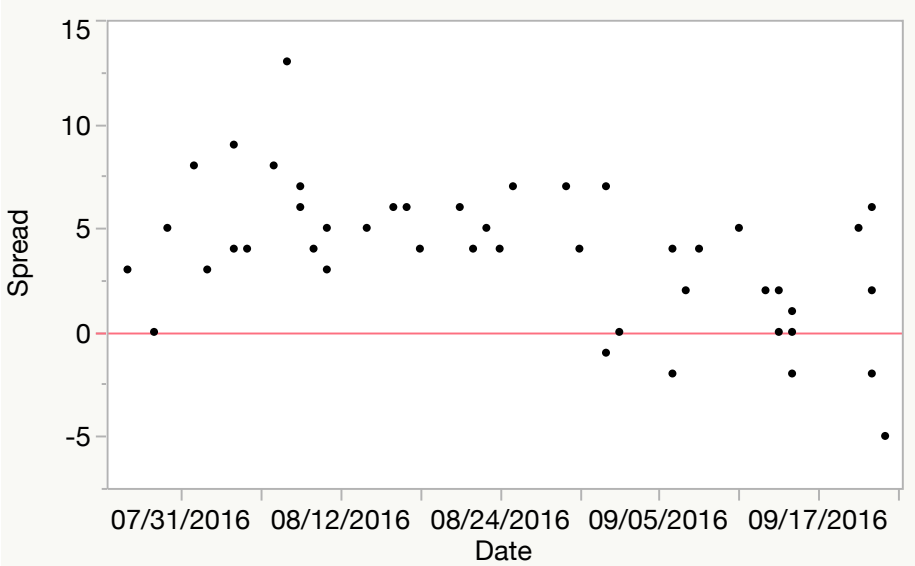
Data collection, survey design/execution

Sampling variation

Example

Surveys prior to recent election





Not just the data, but what you say about data.

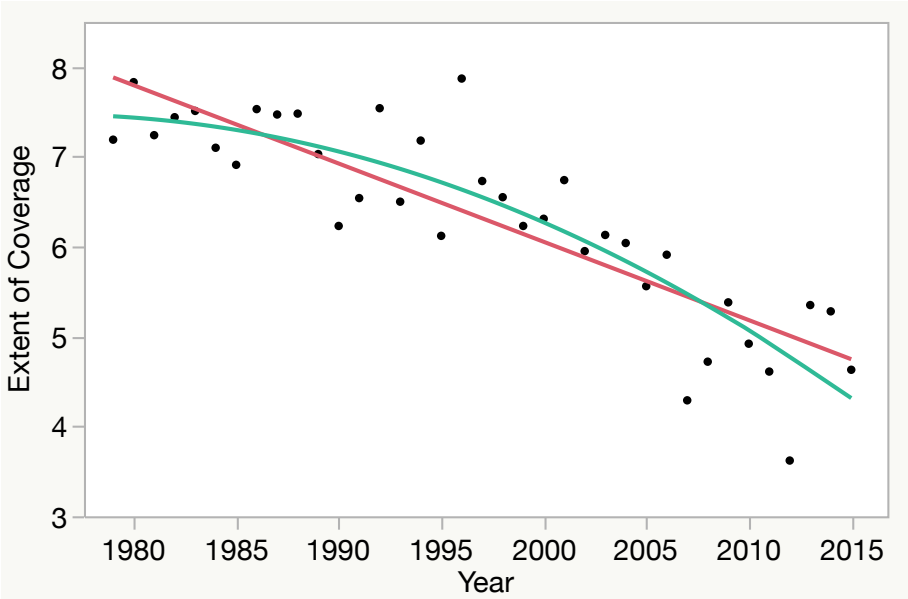
Climate Change

Topic

Polynomial trends, time series

Example

Arctic ice extent (million km^2)



Linear or Quadratic?

Assumptions?

Again how you frame data matters.

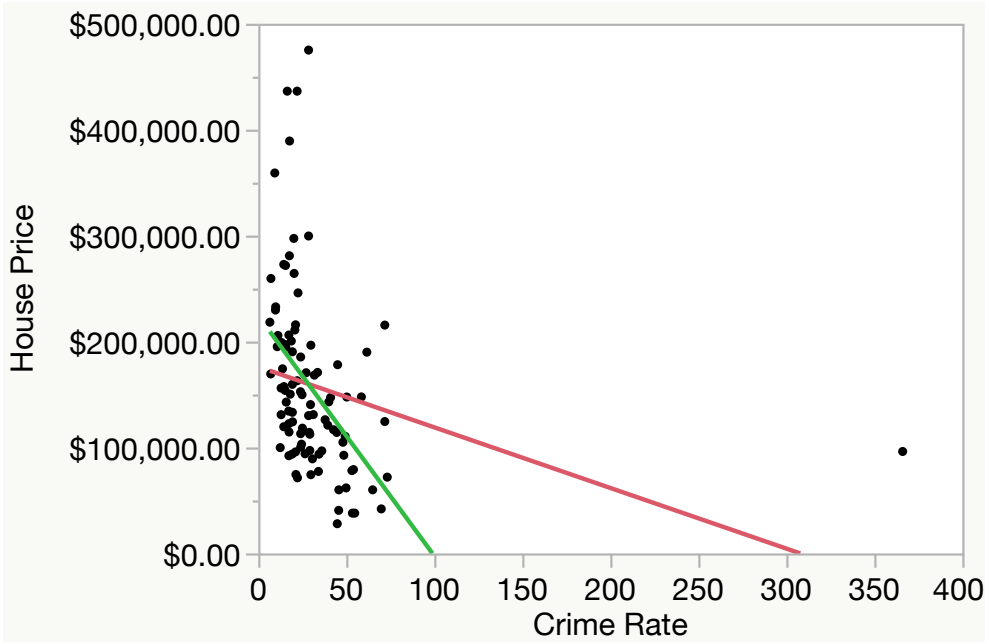
Crime Rates

Topic

Leveraged outliers

Example

Association between crime rates and housing values in communities in Philadelphia area



Where is that?
Why is it so different?

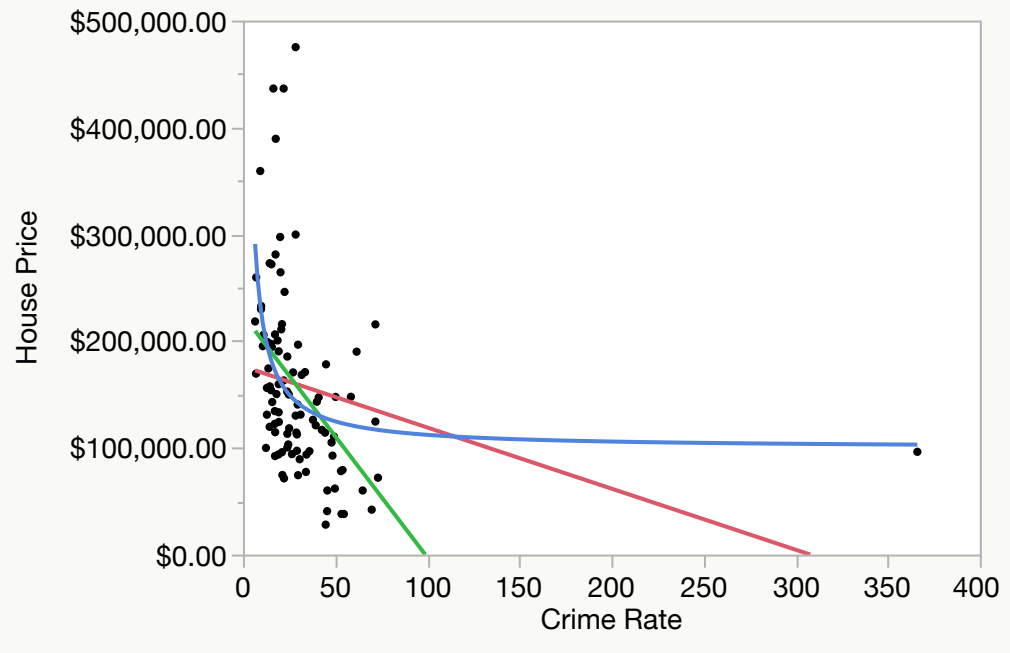
Crime Rates

Topic

Leveraged outliers, nonlinear trends

Example

Association between crime rates and housing values in communities in Philadelphia area



Where is that?
Why is it so different?

Linear?

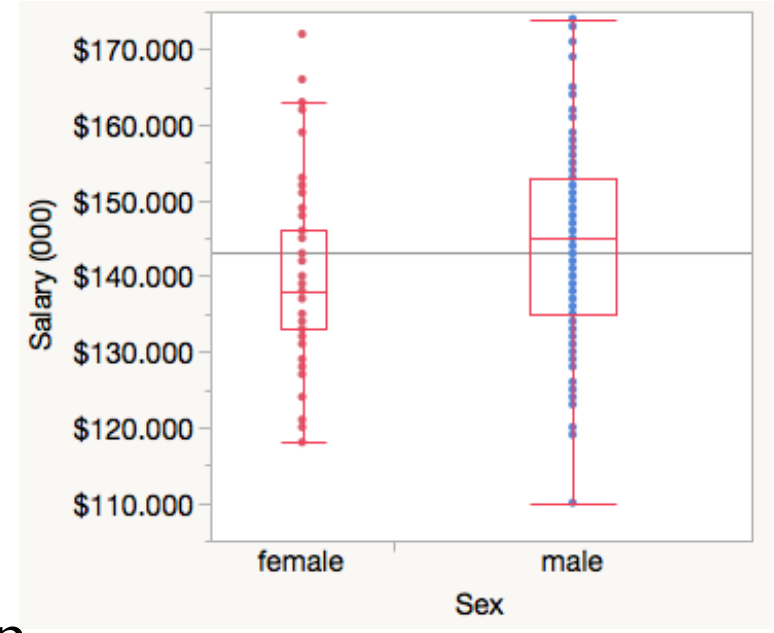
Wage Discrimination

Topic

Dummy variables in regression

Example

Adjusting for confounding effects in two-sample comparisons



male-female
Assuming unequal variances

Difference	4.67045	t Ratio	2.390108
Std Err Dif	1.95407	DF	122.4104
Upper CL Dif	8.53861	Prob > |t|	0.0184*
Lower CL Dif	0.80229	Prob > t	0.0092*
Confidence	0.95	Prob < t	0.9908

Wage Discrimination

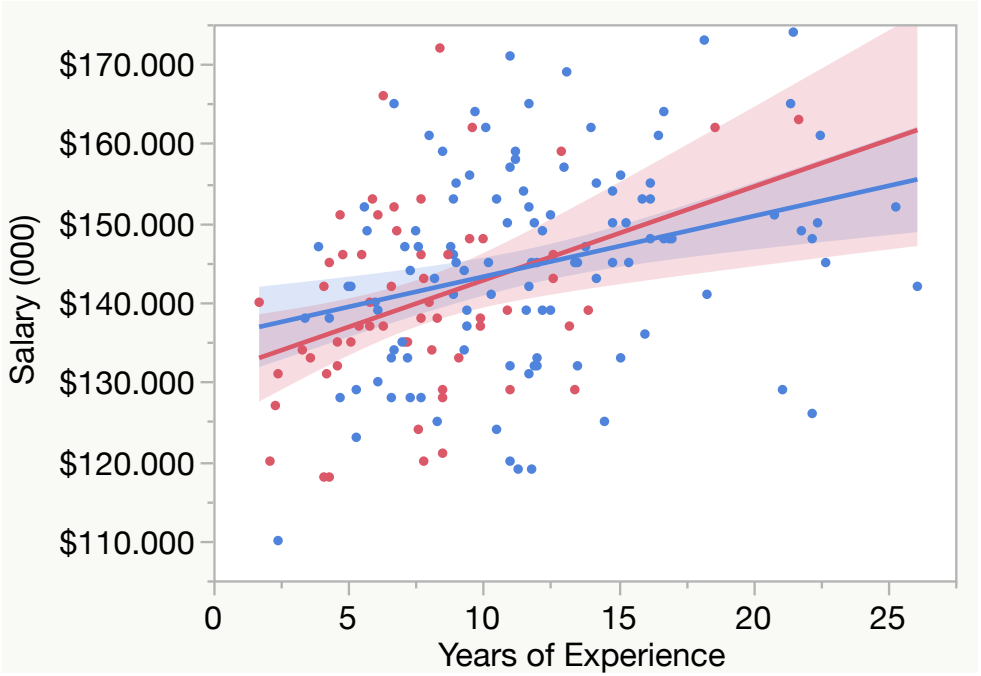
Topic

Dummy variables in regression

Example

Adjusting for confounding effects in two-sample comparisons

Knowing how to use dummy variable is irrelevant if don't know when!

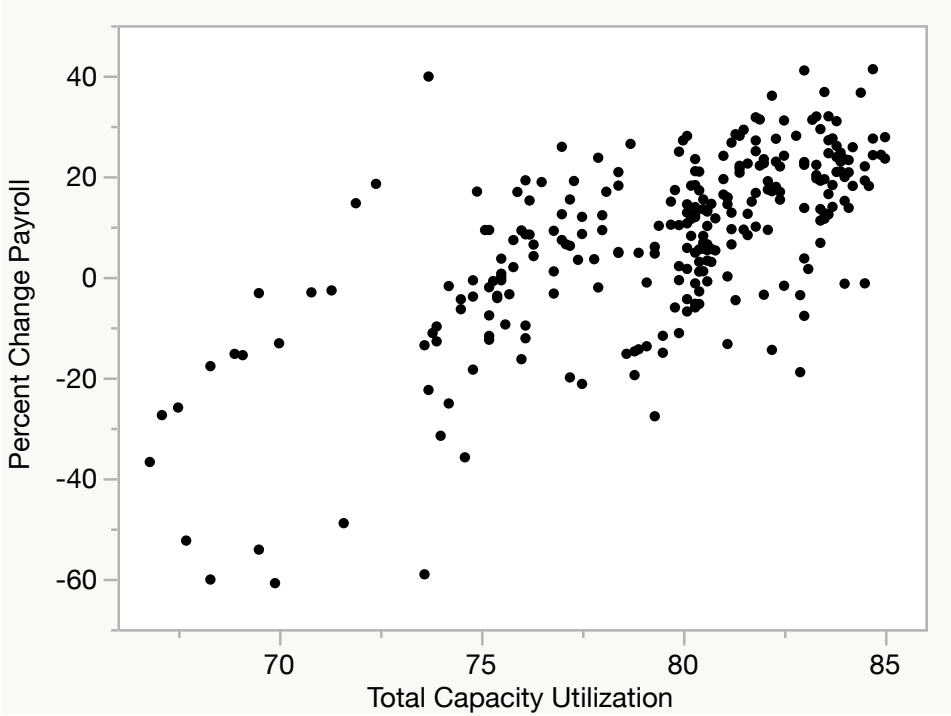


Macroeconomic Data

Macro data common in other courses

Why are there circular patterns in this plot?

- Monthly macroeconomic data
- Payrolls and capacity utilization

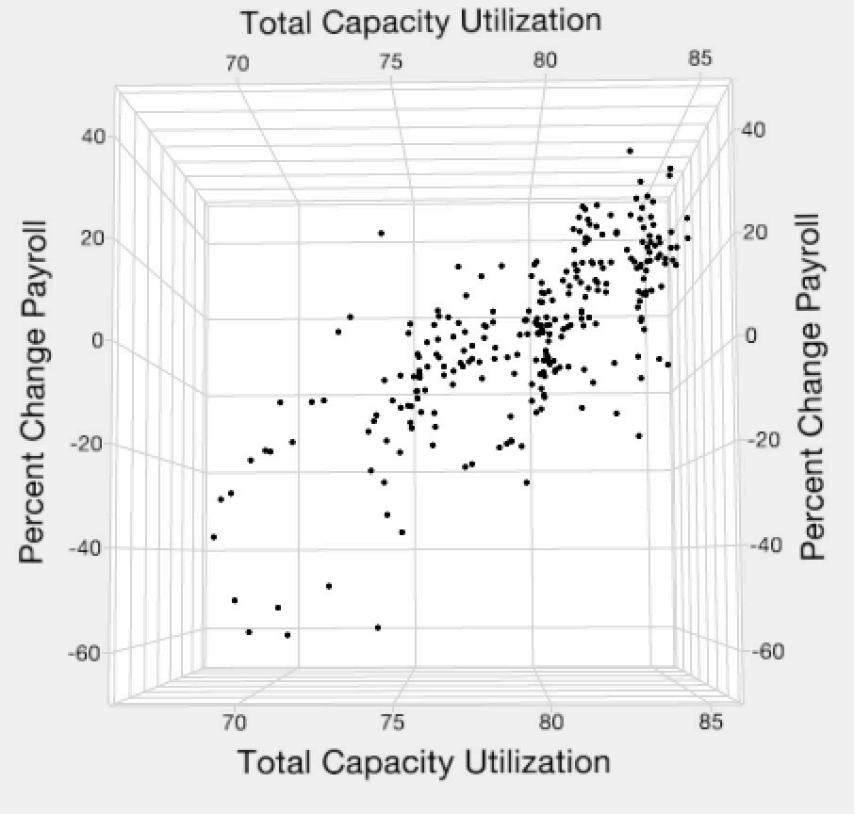


Macroeconomic Data

Macro data common in other courses

Why are there circular patterns in this plot?

- Monthly macroeconomic data
- Payrolls and capacity utilization



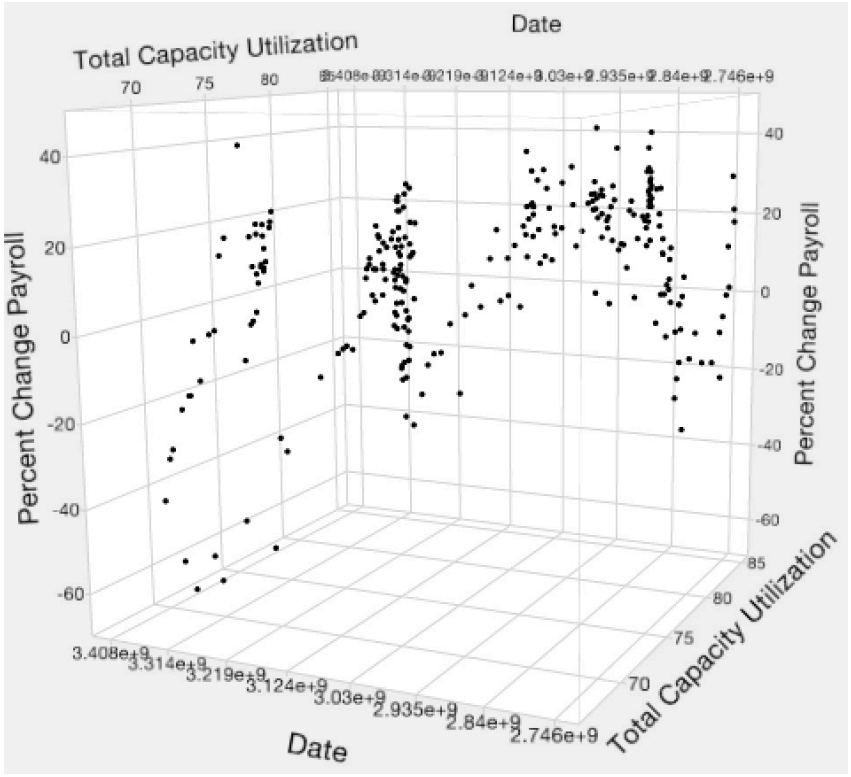
Macroeconomic Data

Macro data common in other courses

Why are there circular patterns in this plot?

- Monthly macroeconomic data
- Payrolls and capacity utilization

At some point to show that Stat actually leverages technology!



Discussion

- Examples draw attention, curiosity
  - Need to be topical to students
  - Have to be fresh
- Collaboration with other fields
  - More important in age of "data science"
  - Reach outside B School: computer science
- Other thoughts
  - Passion
  - Respect