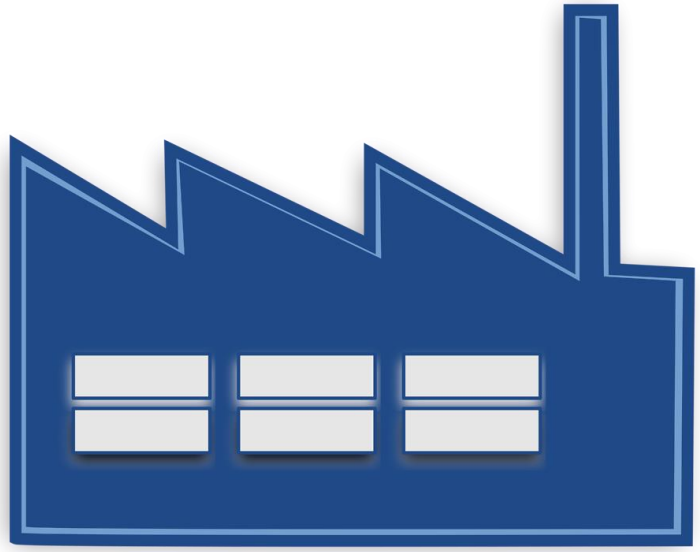
# A Brief Introduction to Statistical Consulting

Mine Dogucu

9.27.2017

# Consulting Environment

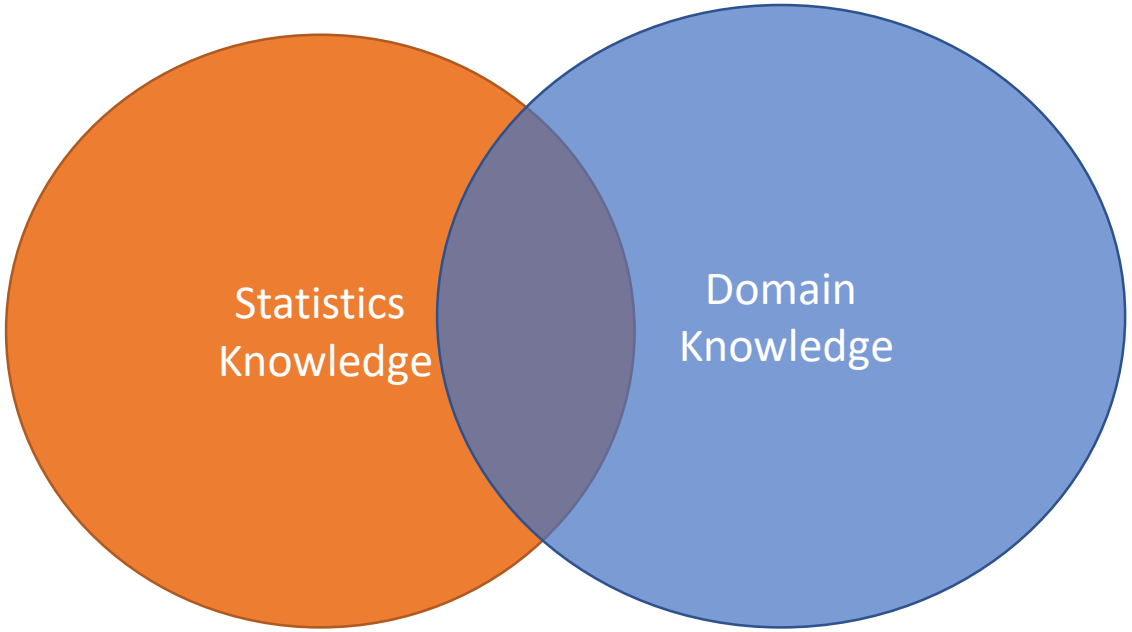
ACADEMIA



INDUSTRY

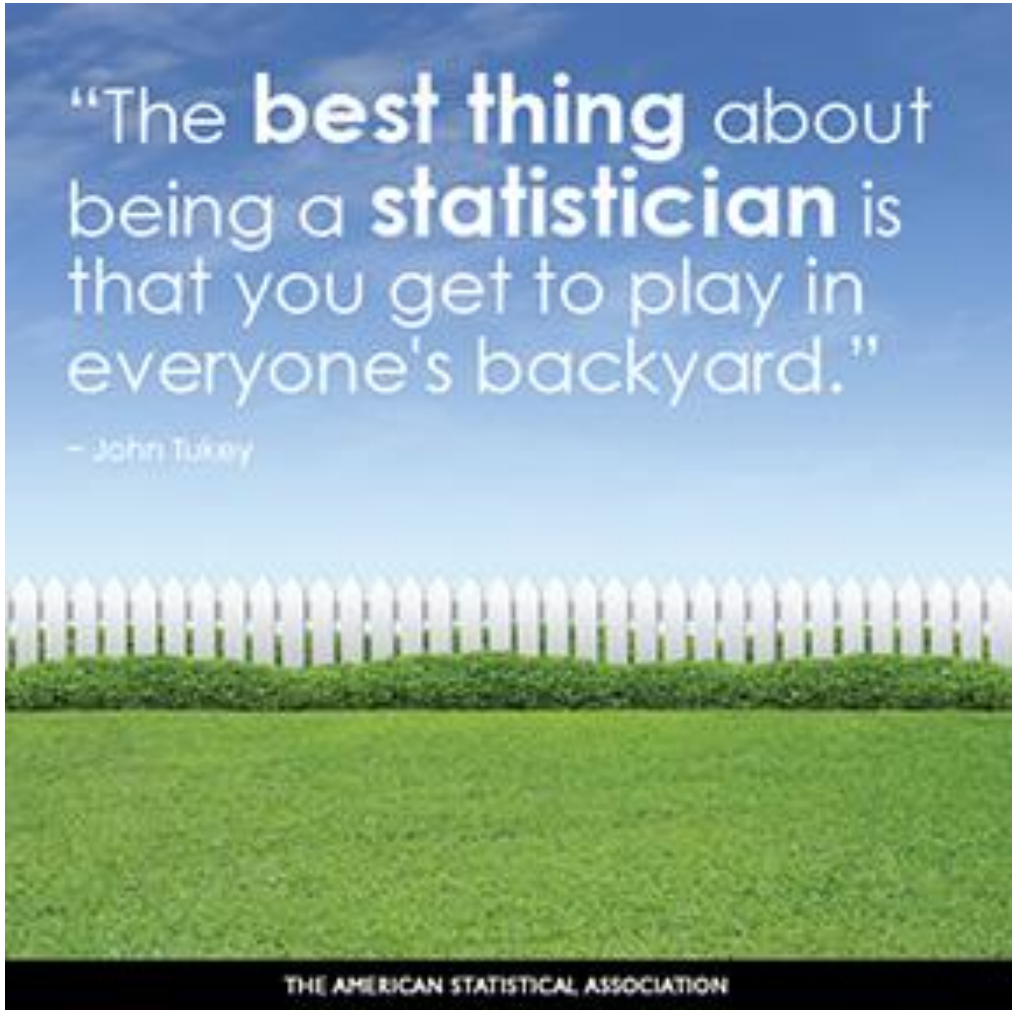
GOVERNMENT

# Domain-Specific Consulting




- In this kind of model, statisticians are expected to have domain-specific knowledge in addition to knowledge of statistics
- This kind of consulting can be in any field but mostly seen in
  - Marketing
  - Finance
  - Pharmaceutical

# Not-Domain-Specific Consulting




- This model is mostly seen in academia and consulting firms.
- There are many fields that rely on external consulting
  - Economics
  - Medicine
  - Public Policy
  - Education
  - Criminology

# Different stages

- Clients can come at different stage(s) of the process.
  - Formulating a question
  - Designing a study
  - Designing an instrument
  - Data collection
  - Data analysis
  - Writing

# Measures of Success

- Because there are different statistical consulting models, there is no single measure of success.
- Thus, it is hard to define rules of effective statistical consulting.
- Some possible measures of success:
  - Whether sales were predicted with reasonable margin of error
  - Whether a candidate wins an election or not was predicted correctly.
  - Whether the number of smokers in a given geographical area is reduced.
  - Whether a grant is funded or not
  - Whether a paper is published or not.

# The Role of Consultant



Give a woman a fish and you feed her a day. Teach a woman to fish and you feed her for a lifetime. Give a woman a good consultant and she can teach herself how to fish.

- The consultant might do all the analyses from A-Z and report the results –more common in industry and government.
- The consultant might help the client do the analysis (capacity building) – more common in academic settings

# The model I worked at

- The RMC - Academic (There are many different models of this as well).
- Around two clients per week (not necessarily new).
- Goal is capacity building
- I almost never did analysis for clients (in very rare cases of high-stakes research).

# Process from A-Z

- Client contacts the RMC
- Client gets a consultant
- Client comes in for in-take meeting (in-take form)
- Client details the problem.
- At the meeting action list is put together.
- After the meeting consultant writes a summary and the action list
- After the meeting both the client and the consultant completes action items.
- Clients cannot return without having completed action items.

# The core is problem solving

- The client has a problem
- The consultant should have a solution
- Example:
  - A study examining family savings behavior
  - Experimental design (control group, phone call, matching more $)
  - The study was run over 6-8 years
  - Paper is submitted and concluded that these treatments do not work
  - The reviewers want to know if the study had enough statistical power.

# Lesson 7: You have to be organized

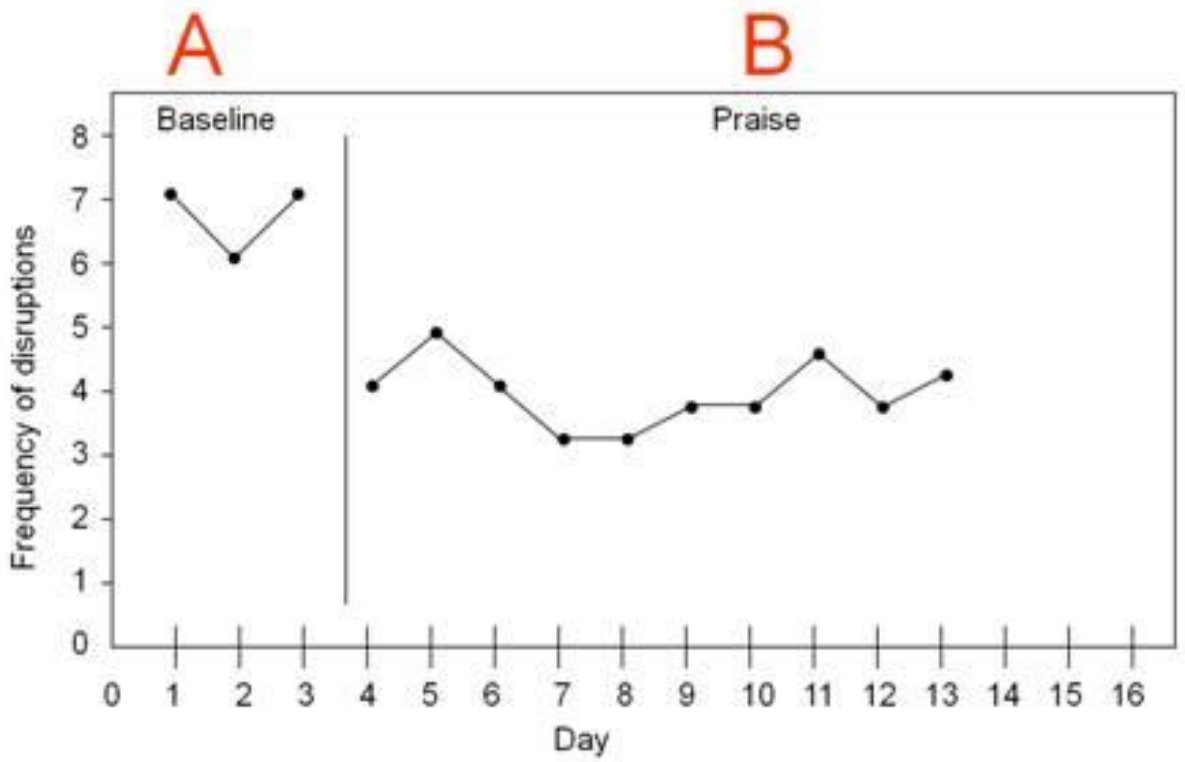
- Your clients probably have deadlines, so do you.
- While working with multiple clients you do not want to mix clients' information.
- Keep track of ANY work you do:
  - So much of statistical work goes undocumented because statisticians are expected to help others.
  - Take a note of ANY statistical help you provide.

# Lesson 6: Some questions are asked over and over and over again

- Keep database.
- Keep resources organized.
- Do not reinvent the wheel.
- Make a blog.

# Lesson 5: There is no way for you to know every statistical model/study design

- Single subject design as an example



# Lesson 4: Communication is as important as statistics itself

- Knowing statistics is not important in consulting unless you can communicate it.
- Three forms of communication in consulting
  - Verbal
  - Written
  - Visual
- You need to communicate using client's language.

# Lesson 3: You need to care about ethics


- Your clients may not necessarily be aware of ethics of research, as a statistician it is your responsibility to make sure you comply to rules of ethical research.
- E.g. Your client may send you an identified dataset that they should not be sending.
- These guidelines are usually are outlined by Institutional Review Boards (IRBs).

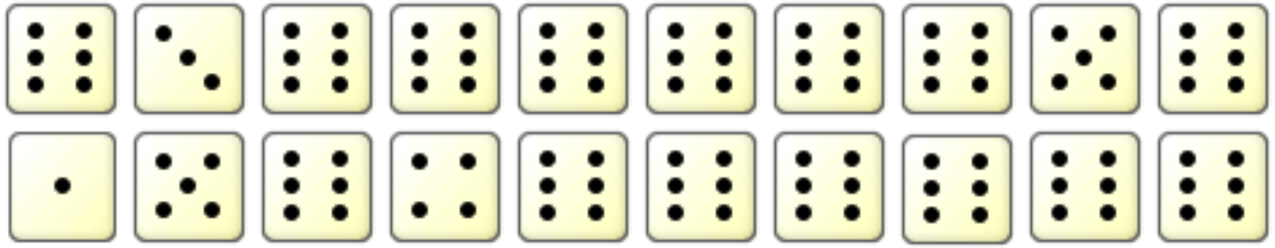
# Lesson 2: There are no magical numbers in statistics

- That includes $\alpha=.05$!
- You are rolling a die in a game in which if you throw a higher value then you win the round. You suspect that your opponent is cheating. You want to test whether they are using a fair or loaded die in this game.
- You decide to observe the next 20 rolls of the die. What is your critical value?

# α=?

- I decided that if my opponent rolls 11 sixes, I would be suspicious.
- P(11 sixes in 20 rolls) = $C(20,11)(\frac{1}{6})^{11}(\frac{5}{6})^{9}$=0.0000897

# P- value



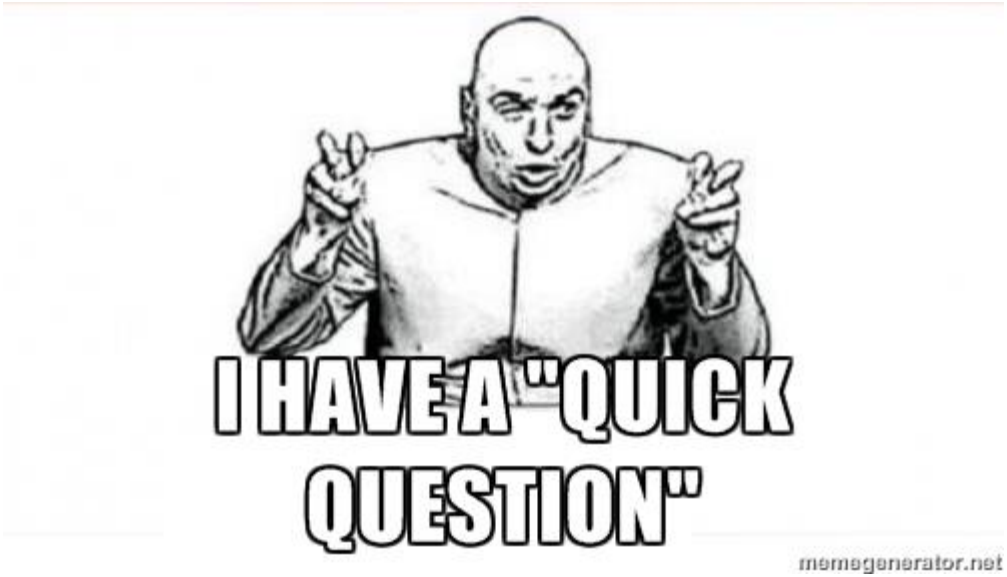
- 15 out of 20 rolls turned out to be 6’s
- P (15 sixes out of 20 rolls) = 0.0000000132516139

# Decision

| Set | | Observed |
|---|---|---|
| Critical value = 11 | < | Observed value = 15 |
| α=0.0000897 | > | p-value = 0.0000000132516139 |

- Just like we can debate how meaningful my critical value (11) is we can also debate how meaningful the commonly accepted(?) α=.05 is.

# Lesson 1: There are no quick questions




Most clients will approach you saying that they have a quick question.

DO NOT BELIEVE THAT!

# THE CORRECT ANSWER TO (ALMOST) EVERY (QUICK) STATISTICAL CONSULTING QUESTION

IT DEPENDS

# Things To Do In Order to Start Your Career Today

- Start and complete your IRB training
- American Statistical Association Membership ($18 until Oct 1 then $25)
  - Section on Statistical Consulting
- Volunteer:
  - Cheikhou Kane
  - Dr. Chris Kottke
  - Dealing with Data II students