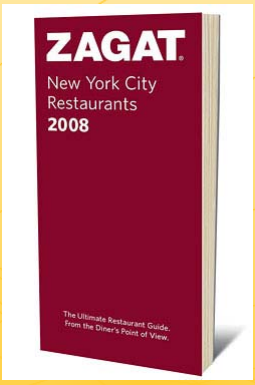
Summary Statistics I

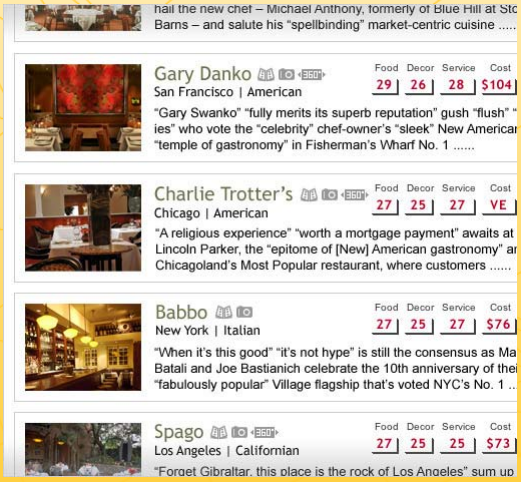
UNIVERSITY OF MINNESOTA

The Story of Zagat



UNIVERSITY OF MINNESOTA

The Zagat Guide …



UNIVERSITY OF MINNESOTA

Computing The Scores

- What should they mean?
  - Popularity
  - Average Rating
  - Probability of You Liking
- How to compute?
  - Frequency
  - Average
  - More Complicated

UNIVERSITY OF MINNESOTA

Secrets Revealed!

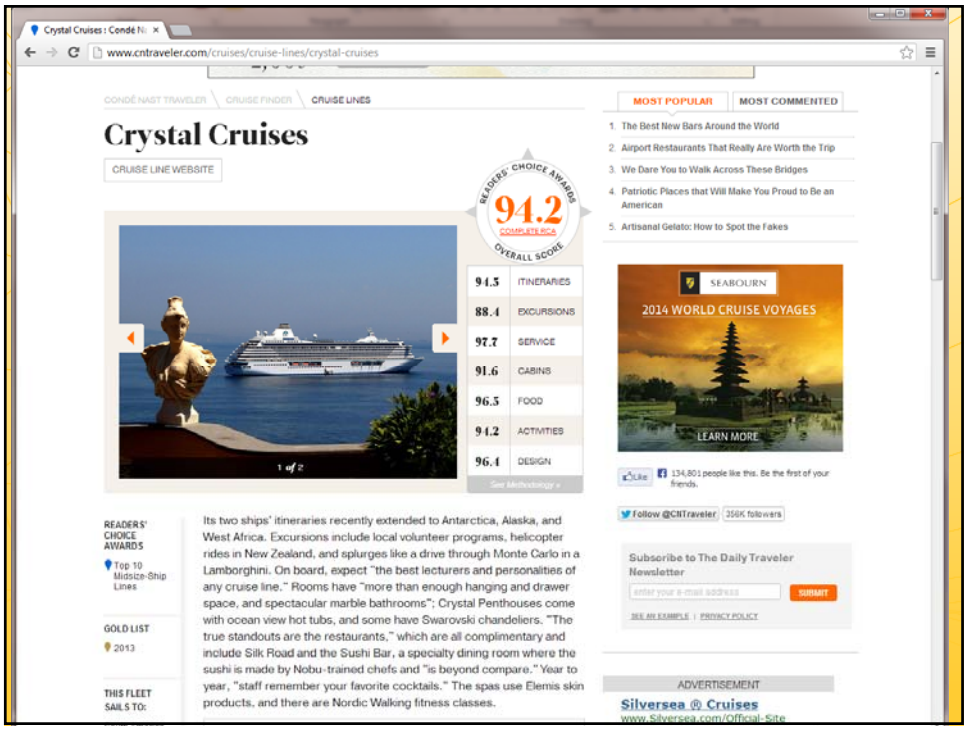
- The “secret” formula

  Rating = {0, 1, 2, 3}

  Score = round (MEAN(ratings) * 10)

  – OK, maybe not so secret – but effective!

UNIVERSITY OF MINNESOTA



Same idea, different formula

- Conde Nast Traveler tallies the percentage of people who rate a particular hotel, cruise, etc. as “very good” or “excellent”
- Relative merits of the two techniques …
  - How do we treat a score of “good” vs. “awful”

UNIVERSITY OF MINNESOTA

Amazon Customer Reviews

UNIVERSITY OF MINNESOTA

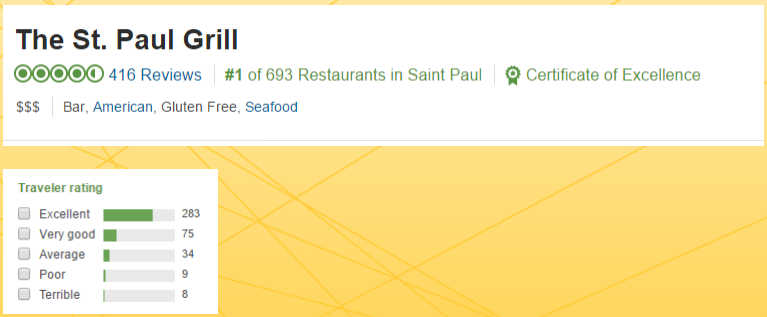
Breaking it Down

- Popularity is an Important Metric
- Averages Can be Misleading
  - Can adjust by summing % who like
  - Can adjust by normalizing user ratings
    - normalization addresses different rating scales
  - May want to consider credibility of individual raters (history of ratings)
- More data is better … to a point
  - Average, Count, Distribution

What's missing here?

- Who you are:
  - If I'm looking for popular new songs, I might not be looking for songs popular among 15-year-old girls
- Your context:
  - If I'm ordering an ice-cream sundae and want a recommendation for a sauce, do I want to hear that ketchup is the most popular sauce?

Tripadvisor.com



UNIVERSITY OF MINNESOTA

Back to Zagat

- Some early Zagat fans argue the guide has been getting worse. Why?
  - Too many mediocre restaurants with good scores
  - Too many excellent restaurants with mediocre scores
- What's happening here?
  - Self-selection bias
  - Increased diversity of raters

UNIVERSITY OF MINNESOTA

Some take-away lessons

- Non-personalized popularity statistics or averages can be effective in the right application
  - Need to understand relationship between average and user need; correct average
- In many cases it can be best to show count, average, and distribution together
- For ranking, one alternative to average is the percentage who score above a threshold
  - Or below!
- Personalization would address many limitations!

UNIVERSITY OF MINNESOTA

Summary Statistics I

UNIVERSITY OF MINNESOTA