What Color is Anonymized Data?

- When it comes to anonymized data, we are colorblind
- Data is anonymized to avoid legal or business constraints on the use/disclosure
- There is no single test or standard for anonymization in the U.S.
  - HIPAA
  - GLBA
  - PUC Regulations
  - Self-imposed
- Tension with technology that continues to make anonymization harder to achieve Business challenge:
  - Find a way to define anonymization OR
  - Protect anonymized information
What’s it Worth?

- Most useful for big picture analysis, trends, patterns
- Usefulness increases as the information approaches the threshold of re-identification: direct marketing, segmentation, targeting
- The tension:
  - Data must remain anonymized but provide maximum utility
  - The incentive is to get as close to the line as possible
  - But where is the line?
- Examples:
  - Payment card transaction data – from economic activity to customer segmentation
  - Anonymized health data worked over with survey data and third party data
  - Aggregate company travel data
- Bottom line: Anonymized data has value inversely proportional to the distance to re-identification
Business, Put on Your Reading Glasses!

- Ask whether service providers will use any data (aggregate, de-identified… magic) for own purposes
- Understand the purposes, risks and benefits
  - Is the data truly anonymized? How can you be sure? Is it worth investing effort to verify anonymization?
  - What is the business getting in return?
  - What is the risk to the business if the data is re-identified?
  - Can the data be mined for confidential company information?
- Typically, there is no reason to allow use of anonymized data outside the limitations of the transaction, unless risk level and cost benefits justify it
- Verify that contract prohibits any use of data, anonymized or otherwise (pay special attention to BAAs, cloud vendors, travel services vendors)
But Wait, What’s on the Other Side?

- Anonymization is a process, not an event, so be creative (within legal and contractual limits)
- If a business is authorized to keep data, keep it
- Examine each step of data processing from the perspective of legal and contractual obligations
- Example:
  - Business is authorized to keep personal data
  - Business keeps a mirror database with UIDs – data may not be deemed de-identified, but not used in this format to provide a product or service
  - When needed, the data is aggregated to create insights