DSC 96
Workshop in Data Science
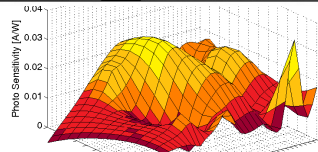
- **Dr. Giorgio Quer**

- **Contact:**
  - gquer@ucsd.edu
  - @GiorgioQuer

- **Office:**
  - SDSC XXX, Mo 6.30-7.30pm
  - Today: write me an email before 6.30pm, meeting online

- **Data Science:**
  - capture
  - maintain
  - process
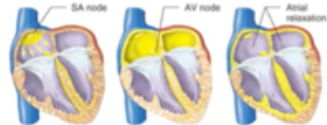 - analyze
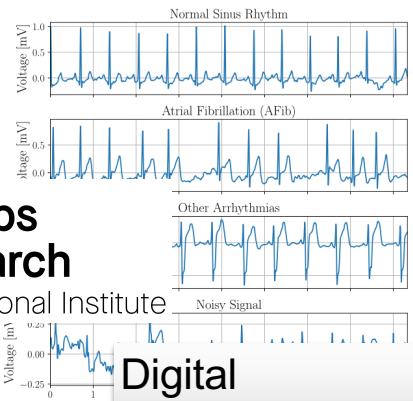 - communicate

From WSN to health

Wireless Networks

Digital Medicine

Remote Sensing

Wearable and e-Health

https://sites.google.com/ucsd.edu/dsc96w2020/
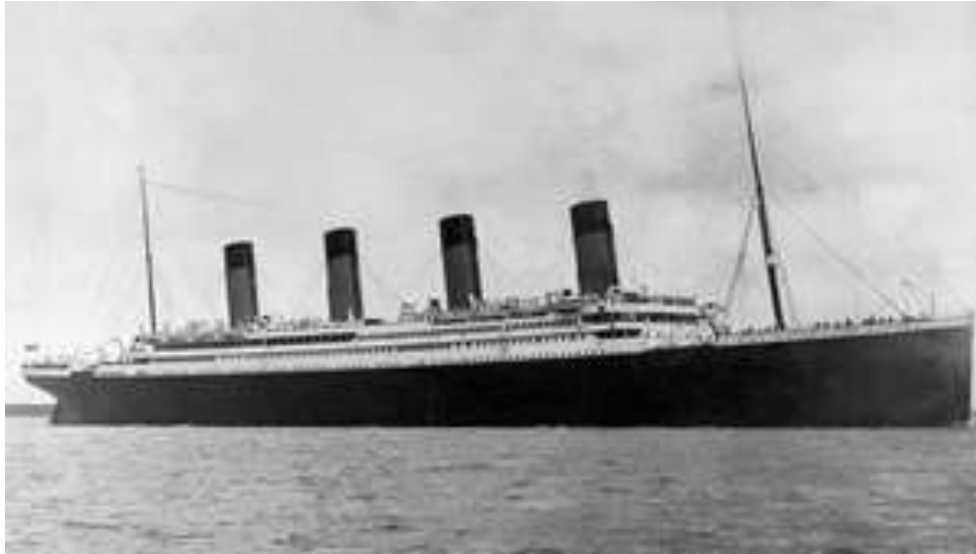
Announcements

- Readings (Week 1 due Friday 1/10)

- Email: gquer@ucsd.edu

- Subject line:

  - [DSC96 W20]: Week 01, Sec A/B, YourFirstName YourLastName

- Email content:

  - Your comments/ questions/ observations on the proposed lectures

- **Tableau basics**

- Intro to interface

- Importing

- Dimensions and Measures

- Chart types

- Finding the story

- Practice on your own

Titanic Data

- How many people were on the ship?
- What was the total of all fares paid?

Tableau

- Open Tableau
  - Create Workbook
  - Import Data: get it from GitHub (link from website, find titanic.csv, click Alt + 'raw')
- Dimensions and Measures
- Chart Types
- Aggregations
- Filters
- Calculated Values

Aggregation Link: https://onlinehelp.tableau.com/current/pro/desktop/en-us/calculations_calculatedfields_aggregate_create.htm

Titanic data

- Pclass:

  - It represent the class for each passenger (first, second, third)

  - 1,2,3

- Tableau

  - It guesses it is a number (like the amount paid): it is in **Measures**

  - It makes sense for Tableau to sum them up or to average them

  - What is the meaning of: average class is 1.85 ?

- Solution

  - Drag pclass it to **Dimensions** !

  - Change the data type to **string**

- What about age?

  - Age in bins (drag to Dimensions, Create, Bins)

  - Fix reasonable intervals (10 years?)

- Titanic dataset and cool things you can do with it: https://www.kaggle.com/c/titanic

- Tableau official training: https://www.tableau.com/learn

- Tableau examples with Titanic data: https://public.tableau.com/search/all/titanic

Titanic questions

- Can you attribute survival to a single primary trait?
- In other words, can you attempt to tease apart the confounding effects of pclass/age/sex, when assessing one?
- Can you explain why the distribution of fares by pclass seems off? why are some 3rd class tickets more expensive than first class?
- Does group size have an effect on survival rate?