Storytelling

Learning Objectives

- How to construct a compelling story?

To start with : What to visualise?

- What to visualise is more important than how to visualise
- In Statistical Analysis : GIGO (Garbage in, Garbage out)
- In Visual Analytics context:

  GIGO → you cannot get much out from inappropriate data being visualised.
- Hence, aim : "Garbage in, Gold out"

Next : Why do you want to visualise?

Two main reasons:
- Understand/Explain Data
  - organise/collate data and explain the trend / features of the data
  - Not to instigate viewers.
  - Letting the data speak.

- Obtain new hypotheses/insight
  - use visualisation to assist getting new hypotheses
  - Let users see the data from different viewpoints.

Then : Set KGI and KPI(s)

Once you have decided the target object/concept of your visualisation :

- KGI (key goal indicator) and KPI (key performance indicator) need to be set.
- KGI will quantitatively evaluate the object/concept to be visualised
- KPI is derived from KGI and describe what needs to be done to achieve KGI.
- e.g, Increase of Sales at EC site
  - KGI : Increase sales,
  - KPI : customer #, price/customer, customer persistence, etc.

Details of KPI

- Example: Web Service
  - One of typical KPI : Customer Persistence
  - How would you decide an appropriate definition of "Customer Persistence"?
  - 5d/week?, everyday?, 1d/month (for expensive items)?

- Does the definition of persistence appropriate for the type of your web service?

- Any other examples?
  - Health related?

Modelling

Different type of variables in modelling

- Objective Variables : a variable to be explained using a model
  - Sales of bottles of water
- Explanatory Variables : variables to explain the objective variable
  - temperature, price, season
- Instrumental/manipulated Variables : Explanatory Variables, whose values can be controlled (changed).
  - price

Selecting variables

- Important to differentiate Instrumental variables from non-instrumental
  - 1 deg increase in temperature yields 7% increase in W.B sales
    - but, you cannot manipulate the temperature…not useful to use this explanatory variables to intentionally increase the sales.
    - price of W.Bs can be manipulated.

- Selecting appropriate Explanatory Variables is important to create and explore the data space through Interactive Visualisation.

Selecting KPIs

Example : EC site analysis

KPI	
Unique Users (UU)	Registered unique users
Daily Active Users (DAU)	unique users logged in per day
Monthly Active Users (MAU)	unique users who used the service at least once a month
Purchased Users (PU)	unique users who purchased
Average Revenue Per User (ARPU)	Sales / DAU
Average Revenue Per Purchased User (ARPPU)	Sales / PU
Persistence	unique users who used the service during specified periods
Conversion Ratio	unique users who managed to participate/purchase certain promotion/target item
Active Ratio	DAU / Total # of members
Withdraw	unique users who withdrew
Clicks	# of advertisement clicks
Page View	access # of certain pages
Cost Per Action	cost to obtain one user (cost of campaign / # of user you obtained)

KPIs should be… - Part I

- Directly related to KGI
  - butterfly effect :
    "Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas? " - Edward Norton Lorenz (1972)

- For the change of value in KPI, unique explanation is associated
  - ARPU (Average Revenue Per User) : Sales / DAU (Daily Active Users)
  - if ARPU decreases, is that because
  - total purchase has decreased?
  - DAU spike due to certain promotion?

KPIs should be… - Part II

- Start with necessity minimum
  - Do not increase the # of KPIs (variables to visualise)

  - A user can understand only small number of visual variables at the time

  - E.g., if "Sales" is KGI, you would only need DAU, ARPPU and Conversion Rate

KPIs should be… - Part III

- Clearly Defined
  - Is "Page View" a good KPI?
  - Page View : increment the count when the page is displayed
  - how about pop-up?
  - how about long page

  - you might need record appropriate user activity logs

  - Instead of Page View, what sort of information would you use?

KPIs should be… - Part IV

- Catergorised based on data types
  - Back to
    - Qualitative
    - Quantitative

    - Categorical
    - Ordinal
    - Interval (you can do +/-)
    - Ratio (you can do +, -, *, /)

Categorise KPI based on Data Types

Example
- USS score (Unit of Study Survey)

(1) Strongly disagree (2) Disagree   (3) Neutral   (4) Agree   (5) Strongly agree

- The average USS score of
  - University was 3.2
  - Faculty of X was 3.4
  - Unit of Study Y was 4.2

➢ What would be the issue(s)?
➢ What would be the appropriate use?
➢ What should we show?

KPI deriving methods

If you don't find KPI, build them!

If you do not have appropriate data for KPI in your original data set, consider deriving them using the existing data by:

- Combine data
- Convert data
- Use Comparative metrics
- Use Summary Statistics

Combining data

Combine multiple data to create a new KPI
- BMI = weight / (height^2)

Converting data

- Metric conversion
  - Fee : $ -> high/middle/low

- To Dimension less data
  - ratio (percentage against the total)
  - Issue? (winning rate, kill-death-ratio) *doesn't tell the size*

- Normalisation
  - Standard score : fit the distribution to (Avg: 50, std: 10)

Using comparative metrics – Part I

Example
- 1st half : Service A had 100,000 UU, Service B had 10,000 UU
- 2nd half: Service A had 110,000 UU, Service B had 15,000 UU

Metrics
- difference: Service A had 10K UU increase, Service B had 5K UU increase
- ratio: Service A had 10% UU increase, Service B had 50% UU increase

➤ Which one is better for measuring impact? *difference* A
➤ Which one is better for measuring growth? *ratio* B

Using comparative metrics – Part II

- Compare against different targets:
  - Many kids use a service "X", hence the service "X" should be customised to suite kids.
  - Should we check the ratio of other age groups?
  - You might find age group >65 might have the similar ratio…. What does it mean?

- Temporal comparison of the same target type:
  - Absolute value
  - Index: Last month's sale was 100, this month's sale was 120, index = 1.2
  - Change: 120 – 100 = 20.
  - Rate of Change : 20 / 100 = 0.2

Using comparative metrics – Part III

- Temporal comparison of the different target types:
  - Union negotiating the wage increase

  - Increase in Company's anual profit   A
  - Increase in wage   B

➤ What KPI would Company use to avoid wage increase?   *wage ↑ higher than profit ↑*
➤ What KPI should Union use to convince wage increase?   *company gets more profit*
                                                                 ⇒ thus wage increase
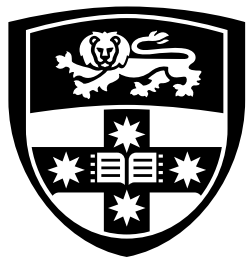
Using Summary Statistics

- Average
- Median
- Most Frequent Value

Summary

- Determine/select KGI/KPI based on your objective

- Understand why you're visualizing certain types of data

- Understand why you're using certain visual variables/methods