Metrics

Overview

- ❏ Simple, Clear, Actionable
- ❏ Audience / market dynamics
- ❏ Technical metrics (load time, crashes)
- ❏ Overall evaluation criteria (OEC)
  - ❏ DAU
  - ❏ Engagement
  - ❏ Watch time, streams, shares

Audience

Consumers

- ❏ DAU, retention, engagement

Creators

- ❏ Paid, region, retention, engagement

Advertisers

- ❏ Reach, perso, CTR

Marketing funnel / Growth funnel

- ❏ Awareness / Aquisition
- ❏ Consideration / Activation
- ❏ Ever purchase / Retention
- ❏ Recent purchase / Revenue
- ❏ Ever recommend / Referal

Product analytics

Lifecycle

- ❏ Initial product ideas
- ❏ Opportunity sizing
- ❏ Experiment design
- ❏ Measurement and launch decisions

Problem types

- ❏ Diagnosing a problem
- ❏ Measuring success
- ❏ Setting goals for a product
- ❏ Launch or not

Product analytics reference for data science interviewing
Created by Nick Jenkins - Data science consultant -
http://www.ndjenkins.com/

Diagnosing a problem framework

- **Clarify**
- **Diagnose**
- **Other products**
- **Segment**
- **Decompose**
- **Summarise**

Clarify

- ❏ High level goal of product
- ❏ Confirm own understanding
- ❏ Audience(s) and use

Diagnose

- ❏ How measured
- ❏ Timing - sudden or stable
- ❏ Internal business issues
  - ❏ Own team or other teams
  - ❏ Technical: Bugs, data, NaN
- ❏ External pressures
  - ❏ Political, special events, natural disasters
  - ❏ Trends, competitors, partners (iOS, TV)
  - ❏ Customer backlash and boycott

Other products

- ❏ Similar changes?
- ❏ No guardrails?
- ❏ Cannibalization?

Segment

- ❏ Age groups, mosaic,
- ❏ Geography, language
- ❏ Frequent or high CLTV customers
- ❏ Premium service customers
- ❏ Platform (mobile vs web)

Decompose the metric

Summarise overall approach

- ❏ Weight towards most reasonable causes
- ❏ Propose how to fix

Setting goals framework

- **High level metric buckets**
- **User actions that support metric**
- **Success metrics, guardrails**
- **Guardrails**
- **Evaluate trade offs**

High level metric buckets

- ❏ Engagement, Retention, Monetization

User actions that support metric

- ❏ How often and long using product
- ❏ Interaction as follows, likes, shares
- ❏ Consider multi-sided (i.e creator uploads)

Success metrics, guardrails

- ❏ Average watch time per user
- ❏ Average # videos watched per session
- ❏ Likes comments subscribe, consider as a funnel

Guardrails

- ❏ Churn
- ❏ Reports
- ❏ Diversity of content

Evaluate trade offs

Statistics

Conditional: P(E | F) = P(E and F) / P(F)

Binomial random variable

- (n choose k) * p^k * (1 - p) * (n - k)
- (a choose b) f(a) / (f(b) * f(a - b))
- $\mu = np$, $\sigma^2 = np(1-p)$

Standard normal distribution: $z = x - \mu / \sigma$

Sampling and confidence intervals

- 1 / n**0.5 rule of thumb
- Increasing sample by 4, spread by 1/2
- Binom SE($\hat{p}$) = ($\hat{p}$ * (1 - $\hat{p}$)) **0.5 / n **0.5

Experiments

Design

- ❏ 50/50 control, test. Paired?
- ❏ By days of week, by geography
- ❏ Random? Stratified?

Units

- ❏ Customers, Businesses, Groups

Power

- ❏ Type II error, correctly rejecting null hypothesis
- ❏ Significance level i.e. 95%
- ❏ Minimum effect - Practical diff
- ❏ 1 / sqrt(n), or 4x sample, 1/2x error

Decisions

- ❏ Link results to goal and OEC
- ❏ Conflicting results
- ❏ Short term vs long term

Problems

- ❏ Multiple test correction
- ❏ Novelty and primacy effect - run or compare only with first time users
- ❏ Group interference - Geo or time based randomisation

Null hypothesis (NH)

- Two samples with metrics; NH they are from same pop, any diff = chance
- Distribution tests (chi-squared, t-score etc.) to produce p-value scores
- P-value = probability of observing the diff or more extreme if the NH is true
- When less than 0.05 we usually reject, but be aware of multiple testing

Power = (1 - b), 80% correctly reject H0

- N = (Za - Zb)**2 / ((μ1 - μ0) / σ)
- Za=2.32, Zb=-1.28, d=4/10, n=81.36