

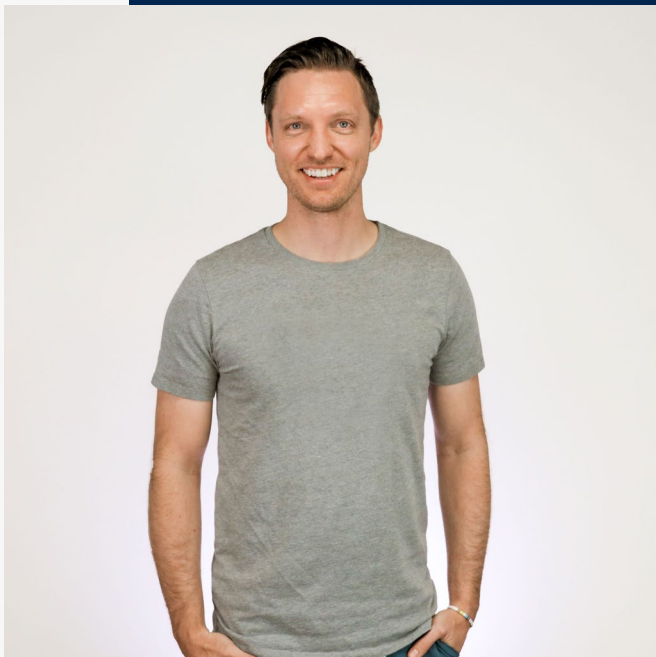
GA4 to BQ  
in 20 Min  
for < \$20



Your Presenter

# Mike Query

CEO / Founding Engineer



- Before I Found Analytics:
  - Dir. of Eng at Ultrasonic R&D (7 years)
  - Embedded systems, signal processing, web apps, data eng
- Founded adMind (2015) to “practice measurement with marketers”
- Last on My Audible: “Make Time”
- If my spouse had the mic, she’d embarrass me by...

## AGENDA

# Today's Agenda

- A Disclaimer
- The Most Common GA4 Data Collection Errors
- How to Spot Them
- How to Fix Them
- How to Keep Them Gone

SECTION

# 01

## BRIEF INTRO TO ADMIND



# We Believe...

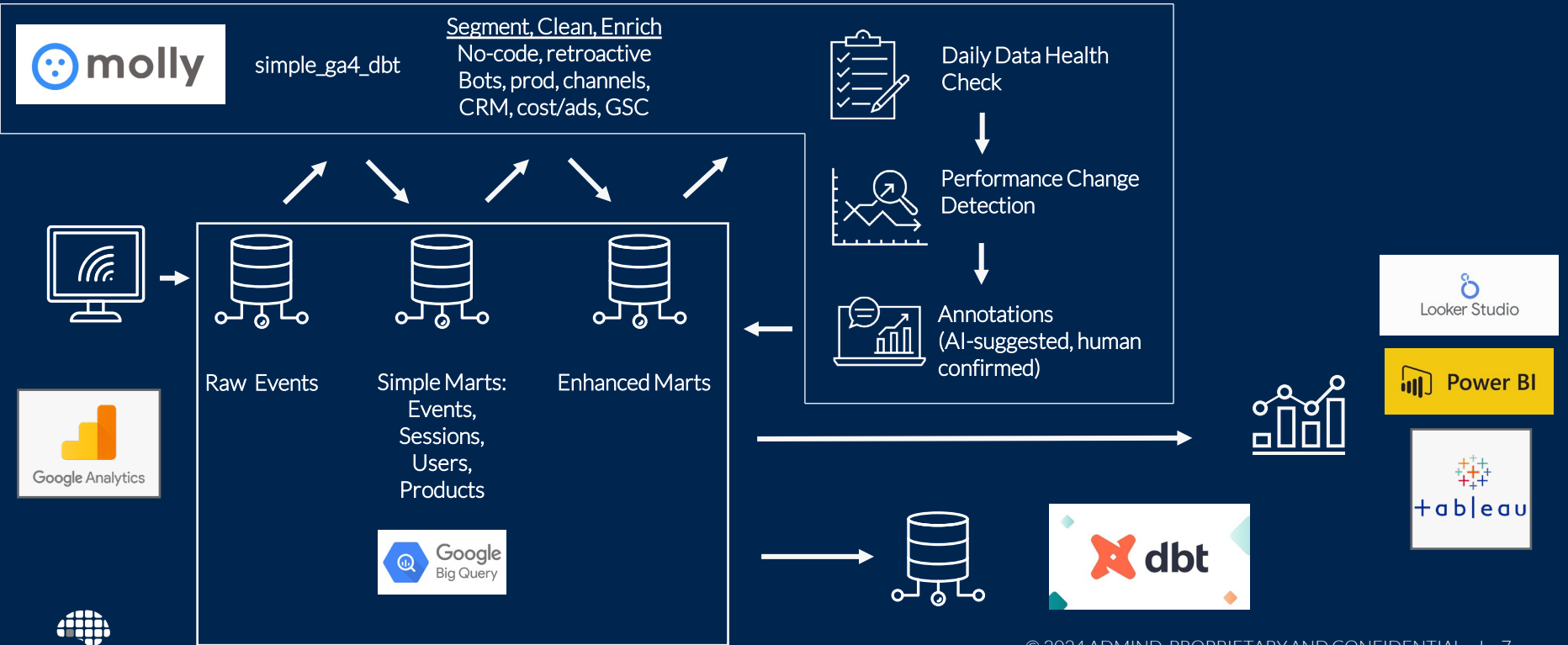
- An insight is “something you don’t know, that you can act on to improve performance”
  - Insights come from analyzing not tracking. Analyze more, track less.
- In rapid, practical AF data collection
  - Sometimes sooo dirty you’d blush telling your mom
  - But..... you gotta wear your seatbelt
- In late data binding. Don’t waste time collecting it if you could just categorize it.
- In warehouses over CDPs
- Dashboards are **products**. Be lean. Iterate. Study the Job to Be Done. If isn’t loved it shouldn’t exist.
- In having your cake and eating it too

INTRO TO ADMIND

# Molly – The Ultimate GA4 Add-On



# Molly – The Ultimate GA4 Add-On



## INTRO

# Join the MeasureCamp Convo!!!

## WHAT'S THERE

- Slack
- Slides
- Repos
- Free Instant GA4 Audit

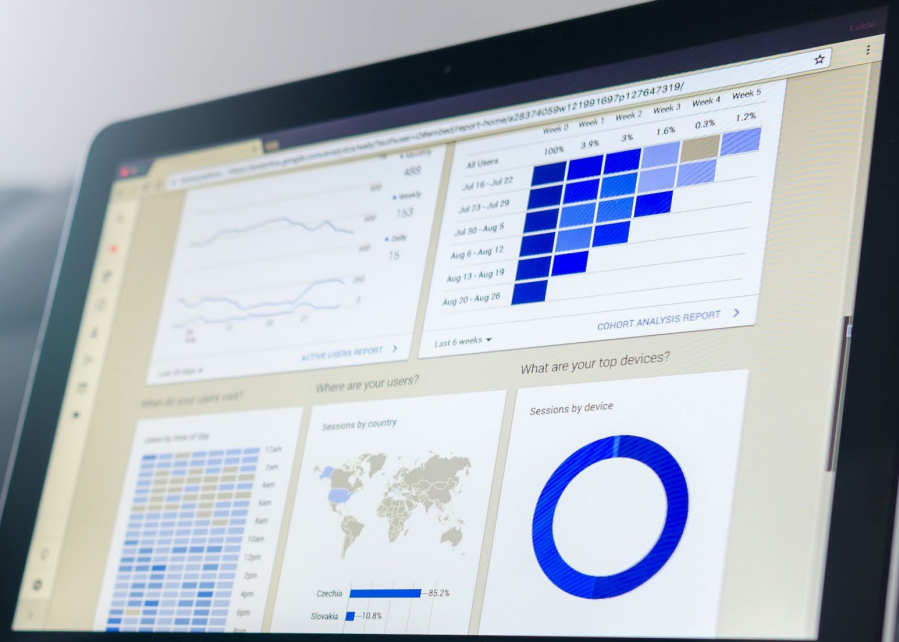




SECTION

# 02

## DOWN TO BUSINESS



TL;DR

# GA4 -> BigQuery Marts with DBT

- Pre-req
  - GitHub acct
  - Google Cloud CLI installed
  - Python 3 (3.10)
  - GA4 to BigQuery connection turned on
- Optional
  - DBT Cloud acct (can be free)
- Pre-req for today



TL;DR

# Clone the Template Repo

- Go To README
- NOTES:
  - mkvirtualenv
  - change values in dbt\_project.yml
  - admindanalytics-ga4
  - profiles.yml



TL;DR

# A Note About Costs

- BigQuery Charges By
  - Storage (\$0.02 / GB / mo – 10GB Free)
  - Compute (\$6.25 / TB – 1 TB Free)
- [Cost Calculator](#)



TL;DR

# Decisions

- Sessions
  - session\_start
- Attribution
  - First, last, last non-direct
  - Prefer paid when gclid regardless of Google processing
  - first\_touch on users, not on sessions... duh
- What to do about path standardization



TL;DR

# Features

- No more (not set) landing pages
- No more (not set) source / medium
- Sharded for query efficiency (match the timezone of account)
- Handle streaming or daily datasets (whichever you have)
- Flatten
- Unique event ids (ooof)
- Fix null ids
- Logic prioritization of the 5 traffic source data sources to consider
- Handle no “session\_start” event sessions
- Attribution
  - First, last, last non-direct (landing pages included)
  - Override when gclid and rationalize contradicting params
  - first\_touch on users, not on sessions... duh
- Prev page, next page, prev prev page, next next page



TL;DR

# Things to Note

- How to run this daily
  - DBT cloud
  - Cron job (ya, annoying I know)
- Model count (simple\_ga4\_dbt vs ga4\_)
- Why is it better (?) to go simple?
  - Fewer models -> lower BQ costs -> lower DBT costs
- Do you need to pre-aggregate to daily totals?



TL;DR

# Things to Know

- Go To README

