

Data Challenge (App Store)

Instructions

- Explore the data included with these instructions and prepare a 45 minute long presentation with recommendations and insights about the App Store
- You may discuss one or more of the suggested topics below in your presentation but they are only suggestions; present whatever you judge to be the most valuable and actionable information from your analysis:
 - Mocking up a business dashboard for tracking the App Store's performance
 - Grouping accounts into meaningfully different segments and analyzing their behavior
 - Grouping apps into meaningfully different segments
- Just like with real data there may be incorrect, suspect, or missing records. If you find a data quality issue please note it in your presentation but analyze the data to the best of your abilities as though there were no issues

Data for this challenge comes with six data tables in CSV format.

- *account_dat.csv* - metadata on accounts
- *app_dat.csv* - metadata on apps
- *in-app_dat.csv* - metadata on in-app content that can be purchased through a given app
- *device_ref.csv* - metadata on devices
- *category_ref.csv* - metadata on app categories (also called genres)
- *transaction_dat.csv* - data on daily downloads and purchases of both apps and in-app content in the App Store

account_dat.csv

- *acct_id* - a unique ID associated with an account
- *create_dt* - date the account was created
- *payment_type* - 'Free' for accounts with no payment method, 'PMOF' for accounts with a payment method on file

app_dat.csv

- *app_name* - the name of the app
- *content_id* - a unique ID associated with a piece of content; all the *content_ids* in this table pertain to apps
- *category_id* - an ID field that denotes the app's category (also called genre)
- *device_id* - the device(s) where the app is available for purchase/download

in-app_dat.csv

- *parent_app_content_id* - the *content_id* of the app where the in-app content can be purchased *content_id* - a unique ID associated with a piece of content; all the *content_ids* in this table pertain to purchasable in-app content
- *type* - the type of in-app content

device_ref.csv

- *device_name* - the name of the device

- device_id - an ID field that denotes the type of device

category_ref.csv

- category_name - the name of the category (also called genre)
- category_id - an ID field that denotes the category of an app

transaction_dat.csv

- create_dt - the date a transaction occurred
- content_id - the content_id of the item that was purchased/downloaded
- acct_id - the id of the account making the purchase/download
- price - the price paid in USD
- device_id - the type of device where the content was downloaded/purchased

INTERNAL reference guide for tables

Data can be found at this Box link:

Internal link:

<https://apple.ent.box.com/s/y8mdd3m382kmp5vatne7bd8r3kplxl7/folder/115951754941>

External link (requires paid Box account):

<https://apple.box.com/s/sw5q7ddt7qmq2y3rjba42f78x56wke0o>

External link (requires iCloud and AppleID account on Apple device):

https://www.icloud.com/icloudrive/0Pn-DIIdbDhKzPWMejYEBW7Og#App_Store_Data_Challenge