CS 327E Project 4, due Thursday, 03/11.

This project makes use of the same shopify dataset from Project 3. If you still have this data on your Jupyter instance, you are good to go. If you don't have the data anymore, you can download by running the following command in a terminal:
`gsutil cp gs://cs327e-open-access/shopify_firestore.zip .`

The goals of this project are to redesign the shopify schema for Firestore, create the database objects according to your model, and populate the objects with the Shopify data.

Using Lucidchart, create an ERD of your Firestore schema. The schema should be modeled based on the design guidelines we discussed in class and the Shopify access patterns we used for Practice Problem 1.

Ensure that your diagram captures the collection of field names and types for each Firestore collection and subcollection. Draw the appropriate relationships between collections. If you're unsure what type of relationship exists between two given collections, consult the Spanner schema for Shopify and/or sample the data. For readability, please use one background color to identify collections and a different color to identify subcollections. Follow the college example from class for other formatting and style guidance. Export your ERD as a `pdf` file and name it `shopify-firestore-erd.pdf`.

Create a new jupyter notebook and name it `project4.ipynb`. Implement the following logic in your `project4.ipynb` notebook:

- Create the Firestore collections and subcollections based on your design and populate them with the shopify data.

- Get a document count for each collection and subcollection you created.

- List the 10 apps with the highest number of reviews (based on `apps.review_count`). Return the `id`, `title`, `developer`, `rating` and `reviews_count` of those apps. Order the results by `reviews_count` in descending order.

CS 327E Project 4 Rubric
Due Date: 03/11/21

Create an ERD for the shopify data in Firestore. Include field names, data types, and ids for the collections and subcollections. Draw proper relationships between collections and subcollections.
      -3 for each missing field name, data type or id
      -3 for each missing or incorrect relationship between entity types
	30

Create a notebook `project4.ipynb` **-3** incorrect file name	3

Create the Firestore database objects that are represented in your ERD
      -5 for each collection which does not match its entity specification
      -4 for each subcollection which does not match its entity specification
      -3 for each field which does not match its entity specification
      -2 for each id which does not match its entity specification
	25

Populate each collection with the appropriate shopify data and get a count of the number of documents per collection and subcollection.
      -5 for each empty collection or subcollection
      -3 for each collection which has missing documents
      -3 for each missing or incorrect count
	30

List the 10 apps with highest number of reviews.
      -3 incorrect or missing filter
      -3 incorrect or missing order by
      -3 incorrect or missing fields in results
      -3 incorrect number of results returned
	12

`shopify-firestore-erd.pdf` pushed to your group's private repo on GitHub. Your project **will not** be graded without this submission.	**Required**
`project4.ipynb` pushed to your group's private repo on GitHub. Your project **will not** be graded without this submission.	**Required**

`submission.json` submitted into Canvas. Your project **will not** be graded without this submission. The file should have the following schema:
 	
{
    "commit-id": "your most recent commit ID from GitHub",
    "project-id": "your project ID from GCP"
}
 	
Example:
 	
{
    "commit-id": "dab96492ac7d906368ac9c7a17cb0dbd670923d9",
    "project-id": "some-project-id"
	Required

}	
Total Credit:	**100**