

Drexel University
College of Computing and Informatics

INFO 323 – Cloud Computing and Big Data

Assignment 1: Ingest Data to Cloud

Due Date: Sunday, April 16, 2023 This assignment counts for 10% of the final grade.
--

A. Requirements

1. Provide full and detailed answer to each question.
2. Work individually. The Drexel University Academic Honesty Rules and Procedures (as stated in the student handbook) will be adhered to strictly.
3. There are 2 questions. The total marks are 100.

B. Questions

Question 1 [20 marks]. Describe and explain the following service models:

- IaaS
- PaaS
- SaaS.

Conduct some research on the Internet; for each type of service model, find some real-world examples and describe them.

Question 2 [80 marks] Hands-on Exercises. Complete the following exercises in GCP. For each step, capture and insert screenshots to show the commands and results.

Step 1. Create a bucket in the GCP cloud storage. Assume the name of the bucket is “YOUR-BUCKET”

Step 2. Open the cloud shell. Git clone the github repository of the textbook “Data Science on Google Cloud Platform” from <https://github.com/GoogleCloudPlatform/data-science-on-gcp.git>

Step 3. Change directory to ‘data-science-on-gcp’. Create a folder ‘data’ under it.

Step 4. Change directory to ‘data-science-on-gcp/data’. Download the monthly data files of the year 2016 “on-time performance data” from the BTC website: https://www.transtats.bts.gov/Fields.asp?gnoyr_VQ=FGJ

Step 5. Display the first two line of the ‘201606.csv’ file.

Step 6. Count the number of lines of the data file for each month.

Step 7. Create sub-directory 'flights/raw' under your cloud storage bucket.

Step 8. Upload the monthly data files of the year 2016 to the folder 'YOUR-BUCKET/flights/raw' in the cloud storage.

Step 9. Ingest the monthly data of the year 2016 to BigQuery in the database 'dsongcp'. Name the table as 'flights-2016'.

Step 10. Run a query on the table 'flights-2016'. List the top-5 origin airports that have the most numbers of flights.

C. What to Hand In

You need to hand in a **Microsoft Word document** including the following items:

1. Your name
2. Course number and title
3. Assignment number
4. Assignment questions and detailed answers to the questions

D. How to Hand In

1. Please name your assignment file as **INFO323-assign1-yourFirstName-yourLastName.docx**.
2. Submit your assignment file through the course website in the **Blackboard Learn** system.

E. When to Hand In

1. Submit your assignment no later than **11:59pm** in the due date.
2. There will be a 10% (absolute value) deduction for each day of lateness, to a maximum of 3 days; assignments will not be accepted beyond that point. Missing work will earn a zero grade.

F. Written Presentation Requirements

Images must be clear and legible. Assignments will be judged on the basis of visual appearance, grammatical correctness, and quality of writing, as well as their contents. Please make sure that the text of your assignments is well-structured, using paragraphs, full sentences, and other features of well-written presentation. Text font size should be at least 11 points.