**Division / Team: Government Digital Services, CC4.0**

**Brief Description of your Div/ Team:** As Singaporeans continue to face an uncertain future, the government has been quick to act to ensure that the country is future ready. One of the most important initiatives is the Smart Nation Programme, which was launched in November 2014.

The programme aims to harness Singapore's infocomm and technology resources to create a more connected and smarter society.

One of the key initiatives of the Smart Nation Programme is Digital Economy and helps to create new jobs and opportunities for Singaporeans.

The Career Coach 4.0 (CC4.0) product is a collaboration between WSG and GovTech which aims to help Singaporeans improve their employability and gain access to job and training opportunities.

The platform also assists coaches to help applicants make informed job decisions and enhance their career journey.

Our role is to build and improve career matching services and ingest data from various sources so that we can find out what both employers and job seekers need and tailor those needs in a unique way.

**Assessment Title:** Take-home Assignment for Data Engineer Intern

**Overview/Background:** This take-home assignment is used to demonstrate your ability to design, build, collect and process data according to a specific set of requirements.

**Assessment Instructions / Case Study Scenarios:**

**Your Task**

**Task 1:**

Steven is a travel blogger that intends to create a travel food series. He is looking at data from Zomato for inspiration. He wants to find restaurants that have good user ratings and interesting past events.

Steven obtained the following data:

· [List of restaurants](https://raw.githubusercontent.com/Papagoat/brain-assessment/main/restaurant_data.json)

· [Country Code (in excel format)](https://github.com/Papagoat/brain-assessment/blob/main/Country-Code.xlsx?raw=true)

1. Extract the following fields and store the data as **restaurants.csv**.

◦ Restaurant Id

◦ Restaurant Name

◦ Country

◦ City

◦ User Rating Votes

◦ User Aggregate Rating (in float)

◦ Cuisines

2. Extract the list of restaurants that have past event in the month of **April 2019** and store the data as **restaurant\_events.csv**.

◦ Event Id

◦ Restaurant Id

◦ Restaurant Name

◦ Photo URL

◦ Event Title

◦ Event Start Date

◦ Event End Date

*Note: Populate empty values with "NA".*

3. From the dataset (**restaurant\_data.json**), determine the threshold for the different rating text based on aggregate rating. Return aggregates for the following ratings only:

◦ Excellent

◦ Very Good

◦ Good

◦ Average

◦ Poor

**Requirements/Expectations**

1. Your code repository should contain a README.md that includes the following:
   * Instructions on how to run your source code locally on our laptop; We need to minimally be able to launch and test your solution locally
2. Preferred language: Python
3. Writing Unit-test is a plus
4. Your code must be hosted on GitHub, or any other similar service, in a publicly accessible repository (e.g., GitHub / Bitbucket / GitLab).
5. You may include a section with the assumptions, interpretations you have made about the requirements above or notes on your architecture decisions.
6. Do show the progress of your work with atomic git commits.

## **Important!**

* We will assess your submission holistically (i.e. not just in terms of functionality), including factors such as:
  + Readability and code cleanliness
  + Secure coding practices
  + Code structure/design, e.g. modularity, testability
  + We will try to run your solution locally. Please ensure that it works, and all required environment variables are stated in the README.md.

**(Optional) Task 2:**

* **CC4 Parking Pte Ltd** wants you to design a Carpark Availability API to allow the mobile application, **ParkingGoWhere** to display empty parking lots based on either the user’s location or user input location.

Using the following data as a reference:  
 [Carpark Details](https://sgc-word-edit.officeapps.live.com/wog/wsg/career-coach-4.0/cc4/cc4-brain-assessment/-/blob/master/system-design/data/HDBCarparkInformation.csv): This provides information about various carparks

(Please refer to the csv file for carpark details)  
 [Carpark Availability API](https://api.data.gov.sg/v1/transport/carpark-availability): This provides the current availability status of carparks

You are to

1. Design the API/s required by the mobile application to output empty parking lots

**Requirements/Expectations**

1. An architecture diagram of the infrastructure/s required to host your API. (Usage of AWS services is preferred)

2. A system design diagram that provides the logical flow of the carpark availability API.

3. A database schema that depicts the tables together with fields required as well as the data type of the fields. This is for both storing and accessing data required by the API/s.

4. An API document that provides the information of the API/s required.   
 (Swagger documentation is a plus)

**Assessment Criteria**

1. Costing, Performance and Scalability
   1. Can the system handle the influx of requests during peak hour traffic?
2. Handling of Error and Edge Cases
   1. E.g. if the user inputs a non-existing location, if the lot is being taken up while a user is travelling to the location etc.
3. API Design
   1. Consistency in naming conventions, data formats and error handling
   2. RESTFul Principles (using of proper HTTP methods etc.)

## **Important!**

1. You **DO NOT NEED TO** develop the API/s or spin up the actual infrastructure.
2. You can document the above in the README.md.
3. Assume that authentication and security is already taken care of.
4. Creativity is a plus. You are welcome to suggest additional features for the application. E.g. “Chope-ing” of parking lots (how to enforce). etc.

**Things to take note:**

* Task 2 is **OPTIONAL!!** Attempt only if you have enough time.
* For questions relating to the assessment, please contact [CHONG\_Qi\_Da@tech.gov.sg](mailto:CHONG_Qi_Da@tech.gov.sg)

**Deadline:**

* Please complete the assessment within **five (5) calendar days** upon receiving this email. Please email your presentation and example of content. Thereafter, you might be invited for a chat with our tech assessors to discuss your submission.
* Please email your presentation and example of content to Nadiyah Yusof (nadiyah\_yusof\_from.pserv@tech.gov.sg). Do title the subject of your email as SWE Career Coach (CC4.0) Internship Tech Test Submission - Your Name (Eg: SWE Career Coach (CC4.0) Internship Tech Test Submission - Carmen Loke).