DJS Compute powered by aws

DJS COMPUTE INTERVIEW TASKLIST

General Instructions:

1. Attempting these tasks is not mandatory, but doing so will give you an advantage
2. The tasks assess your basic understanding and knowledge. Feel free to use online resources to learn and perform the tasks accordingly.
3. Attempt as much as you know; it is not mandatory to complete all the tasks.
4. You can perform tasks of multiple domains.
5. Do not copy tasks just for the sake of completion.
6. Create a GitHub repo named `djs-compute-tasks` and upload the jupyter notebook/colab/code-base there.
7. Provide the Github link to your task in the registration form
8. For doubts, contact the mentors listed under departments or contact CP and VCP from the contacts below.

- Aaditya Malani - +91 99876 79904
- Fiona Haria - +91 82916 78256

WEB DEVELOPMENT TASK:

Task 1:

Imagine you're building a digital calling card to introduce yourself as a web developer! Create a one-page portfolio website using either HTML/CSS or React. Showcase your creativity: Present yourself with an "About Me" section, highlight 2-3 projects (school or personal) with descriptions and screenshots, and include a way to contact you (just your email for now).

You can make these pages:

About Me: A section with a brief introduction about yourself, your skills, and your goals as a web developer. Add a link to your resume as well.

Projects: A section showcasing 2-3 projects you've worked on (school projects or personal projects are fine). Include a short description, links (if applicable), and screenshots/images for each project. (If you don't have any projects, put dummy projects)

Contact: A dummy contact form which serves as a medium to contact you.

Feel free to add any other pages as well. To add new pages, you can refer routing in react.

🌟**Brownie Points:**

- Pushed code on github
- Responsive Design
- UI/UX
- Use of Bootstrap/Tailwind
- Animations
- Additional Libraries
- Hosted website

Task 2:

Dr. Sarah Jones specializes in pediatrics and urgently needs a user-friendly website for appointment booking. Her current system relies on phone calls, leading to frequent misunderstandings and missed appointments.

Create a website on ReactJS. The new website should cater to both tech-savvy and less-familiar parents. Here's what Dr. Jones envisions:

- **Home page** for showing doctor's information, contact details and address on any preferable map.
- **Booking functionality** in a form with necessary fields and with validation such that date should be in the future.
- After booking there should be a page to display the booked appointment with all the details.

You don't have to code the backend, just code the frontend.

Figma template for reference only:
https://www.figma.com/design/2qwIwQkOaoVGcdnHTsYoV5/Appointment-Booking?node-id=2-2&t=Pq9DJVaxwGSab78T-0

Don't copy this, also feel free to add multiple pages

🌟**Brownie Points:**

- The appointment details are accessible even after page is reopened in "past booked appointment"
- Pushed code on github
- Responsive
- UI/UX
- Bootstrap/Tailwind
- Animations
- Additional Libraries
- Hosted website

Resources:

- React Tutorial: https://youtu.be/f55qeKGgB_M?si=CUx9EvKMySfqMpjP
- Github Tutorial: https://youtu.be/HkdAHXoRtos?t=42

Feel free to contact any of the mentors below for any queries:

- Taher Afsar - 8356994540
- Vedant Naik - 7506058458
- Kartikeya Pandey - 9321791268
- Krish Thakkar - 7219776463

DATA ANALYTICS AND MACHINE LEARNING TASKS

Task 1: Basics of Python

This task is designed to test your basic understanding and knowledge of programming in general and Python in specific. It covers data types, operators, handling user-defined input, strings, string methods, lists, list comprehension, sets, tuples, dictionaries, basic object-oriented programming concepts and regular expressions. We strongly advise that you brush up on these concepts before coming to the interview.

Beginners who are unfamiliar with Python are advised to watch the following tutorial before beginning this task.

Python Tutorial: https://youtu.be/__uQrJ0TkZlc?feature=shared

Task Link:
https://colab.research.google.com/drive/1yT8vjpqAoVIPTGbIaxQQF6oIv2wCma71?usp=sharing

Note: Please make a copy of this Colab notebook in your own drive and solve the questions in the copy. **Do not edit the original file.**

TASK 2 – Data Pre-processing & Visualization

This is designed to see how well you handle raw, unprocessed data (cleaning & transformation), how effectively you can visualize this data and what kind of insights can you derive from these visualizations.

Use Python Libraries such as NumPy, Pandas, Matplotlib & Seaborn to clean and transform the dataset given below. Ensure that you handle any missing or duplicated values, outliers and inconsistent data types. Explore the dataset and use appropriate plots to visualize the data and clearly and briefly explain what you interpret from them. Maintain a README.md file with explanation of your insights and conclusions.

The use of comments to explain what a particular piece of code is doing is highly encouraged.

Dataset Description: The dataset provided includes the players' data from FIFA 21.

Dataset:

https://drive.google.com/drive/folders/1OJRPqvTBzs0YRpJ2LlkiXxrSZ7V_3uY9?usp=sharing

Resources:

1. https://youtu.be/QiqZliDXCCg?feature=shared
2. https://youtu.be/_Eb0utIRdkw?feature=shared
3. https://youtu.be/xi0vhXFPegw?feature=shared

TASK 3 – Model Fitting (OPTIONAL)

Please note that this task is entirely optional but attempting it will earn you **brownie points**.

It involves fitting any suitable Machine Learning model to the given data. Please mention your reasoning for choosing the model. The data may require pre-processing before a model can be fitted to it. Evaluating the performance of your model will be a plus point.

Dataset Description: The dataset given below contains census data. Your job is to apply a suitable model to predict whether income exceeds $50K/year or not.

Dataset Link:
https://drive.google.com/file/d/1MXrmRTJVmMa-RVzQYKqODFKbNRb-aLsq/view?usp=sharing

Resource: https://youtu.be/29ZQ3TDGgRQ?si=u5tY-QS9Cn7tP-8R

Feel free to contact any of the mentors below for any queries:

- Ananya Godse - + 91 7506090969
- Aryan Chintakindi - +91 9082503502
- Vaishnavi Shridhar - +91 9930608762
- Karthik Nambiar - +91 7045210259

DATA ENGINEERING TASKS

Task 1: Python

The tasks of data analytics and machine learning are to be followed. Please refer to page 4.

Task 2: SQL

Resources:
Video: https://www.khanacademy.org/computing/computer-programming/sql
Reading: https://roadmap.sh/sql
Setup

a) Windows:
How to install MySQL 8.0.37 Server and Workbench latest version o...
b) Mac: How to Install MySQL on Mac | Install MySQL on macOS (2024)

Task Description:

1] Database and Table Creation:

Create a new database named Social-Media.

A] Users table:

- user_id (Primary Key): U001, U002, U003...
- username: social_butterfly, tech_guru, fitness_fanatic...
- email: user001@email.com, techie@email.com...
- join_date: 2023-01-15, 2022-11-30, 2024-03-22...
- country: USA, India, Brazil...
- bio: "Love to travel and meet new people", "Tech enthusiast and coder"...

B] Posts table:

- post_id (Primary Key): P0001, P0002, P0003...
- user_id (Foreign Key referencing Users): U001, U002, U003...
- content: "Having a great time at the beach!", "Just launched my new app!"...
- timestamp: 2024-07-14 15:30:00, 2024-07-14 16:45:00...
- likes_count: 42, 78, 15...

- comments_count: 5, 12, 3...

C] Comments table:

- comment_id (Primary Key): C0001, C0002, C0003...
- post_id (Foreign Key referencing Posts): P0001, P0002, P0003...
- user_id (Foreign Key referencing Users): U002, U003, U001...
- content: "Looks amazing!", "Congrats on the launch!"...
- timestamp: 2024-07-14 15:35:00, 2024-07-14 16:50:00...

D] Friendships table:

- friendship_id (Primary Key): F0001, F0002, F0003...
- user_id1 (Foreign Key referencing Users): U001, U002, U003...
- user_id2 (Foreign Key referencing Users): U002, U003, U001...
- status: "accepted", "pending", "rejected"...
- created_at: 2024-05-20, 2024-06-15, 2024-07-01...

E] UserInterests table:

- interest_id (Primary Key): I0001, I0002, I0003...
- user_id (Foreign Key referencing Users): U001, U002, U003...
- interest: "Travel", "Technology", "Fitness"...

2] SQL Queries:

1. List all users who have posted more than 5 times, along with their post count.
2. Find the top 3 countries with the most users, and for each country, show the most common interest among its users.
3. Find users who have mutual friendships (both have accepted each other's friend requests) and share at least one common interest.
4. List users who have commented on their posts, along with the post content and comment content.
5. Create a stored procedure that takes a user_id as input and returns that user's posts, along with the average likes and comments for each post, compared to the overall average likes and comments across all posts.

3] Git Integration:

- Initialize a Git repository in your project folder.
- Create a branch named sql-task.
- Commit your SQL scripts with appropriate commit messages.
- Push the changes to your GitHub repository and provide the link.

🌟**Brownie Points:**

Bonus Task 1: Implement a trigger that automatically updates the comments_count in the Posts table whenever a new comment is inserted or deleted.

Bonus Task 2: Find the "influence score" for each user, calculated as the sum of their post likes plus the count of their accepted friendships, but only considering friends who have posted in the last 30 days. List the top 10 users by this score. (Joins, Aggregation, Subquery

Task 3: Spark

Data Engineering with Apple Stock Prices

Resources:

- Video: Setting up the PySpark environment on Google Colab
- Reading: https://www.scaler.com/blog/data-engineer-roadmap/
- Platform: Google colab
- Task Dataset: https://www.kaggle.com/datasets/fredrickkariuki/aaplcsv

Task Description: Utilize your data engineering skills to process and analyze historical Apple stock prices. This task will test your ability to clean and transform data using Apache Spark with Scala or Python(Pyspark), preparing the dataset for insightful analysis.

1] Dataset Description:

The dataset consists of historical Apple stock prices with the following columns:

- Date: Date of the stock price
- Open: Opening price

- High: Highest price of the day
- Low: Lowest price of the day
- Close: Closing price
- Adj Close: Adjusted closing price ● Volume: Volume of shares traded

2] Task Breakdown:

1. Initializing a Spark Session

   - Write a Scala/Python(PySpark) script to initialize a Spark session using Spark.

2. Reading Data from the CSV File

   - Read the provided CSV file into a DataFrame.

3. Cleaning the Data

   - Fix missing values in the dataset and remove any duplicate rows.

4. Performing Map Transformations

   - Apply map transformations to the DataFrame.
   - Include at least one transformation and one action.

5. Saving the Cleaned Data

   - Save the cleaned DataFrame to a new CSV file.

3] Git Integration:

- Initialize a Git repository in your project folder.
- Create a branch named spark-task.
- Commit your SQL scripts with appropriate commit messages.
- Push the changes to your GitHub repository and provide the link.

🌟**Brownie Points:**

Bonus Task 1: Implement a custom function to clean the data further, such as normalizing the Volume column.

Bonus Task 2: Create a simple summary report of the dataset after cleaning, including statistics like count, mean, min, and max for numerical columns.

Bonus Task 3: Perform a groupBy operation on a derived column like Month (extracted from Date) and calculate aggregate statistics like average Close price and total Volume.

Feel free to contact any of the mentors below for any queries:

- Vanshika: +91 9833200052
- Bhoomika: +91 9967234049
- Harshvardhan: +91 9372636335
- Aayushi: +91 8108143536

MARKETING TASKS

Prepare a database of up to 20 companies, each with their point of contact (POC) name and email address. This task will showcase your market research skills. Utilize online databases, company websites, LinkedIn, and professional networks to identify potential connections.

Prepare a Google spreadsheet with the following columns:
1.Company Name
2.POC Name
3.Email Address

Additionally, having a basic understanding of the Fiverr app can add brownie points for you before the interview. Fiverr is a popular freelancing platform where clients can hire freelancers for various services.

Feel free to contact any of the mentors below for any queries

- Parth Patel - 7506268263
- Delanie Rodrigues - 9321649960
- Arya Mangaonkar - 79771 22658

PROJECT ASSOCIATE TASK

Task 2:

Your client wants an AI-based personal budget tracker that enables users to efficiently manage their finances by analyzing spending patterns and offering budget recommendations.

- **Innovative Approach**: The solution should go beyond traditional budget tracking methods, offering personalized insights and recommendations
- **Thorough Research**: Conduct an in-depth analysis of existing vulnerabilities and user behaviours to inform system design.
- **Technology Stack:** Employ technologies for backend, frontend, and machine learning, ensuring ease of scalability and integration.

The time duration for the project is **3 months**.

Your tasks are:

1. **Define project phases:** Research, Design, Implementation, Testing, Deployment.

2. **Outline specific tasks** for each phase, assigning responsibilities and setting milestones for each department.
3. **Establish timelines** for each phase and milestone, ensuring alignment with overall project objectives.

Create a Word document or PowerPoint presentation for submission, including all reports, proposals, and design documents.

Resources

1. https://www.geeksforgeeks.org/phases-project-management-processes/
2. https://youtu.be/ZWmXi3TW1yA?si=m1uv0uTHKbC9Cwwz

Task 2:

The second task is the same as that of the marketing tasks given above

---

For any queries contact, any of the leads given below

- Vipul Mhatre - + 91 8591592184
- Anshuvi Shah- + 91 75069 07585
- Isha Solanki- + 91 92229 15939

CREATIVES TASKS

Task 1:

Create a creative post grid for an event/seminar or on any festival

Alternatively, You can showcase any previous work done in Canva or in Figma

Task 2:

Create a Figma design for a website/app of your choice
Leo Thompson, a busy CEO in his early 50s from New York, needs a travel booking website that offers quick search and booking, personalized recommendations, detailed itineraries, and 24/7 customer support. He appreciates easy and concise navigation to important parts of the website.

Requirements:

- Header with clear navigation and profile in right
- Featured destinations/deals/categories sections
- Footer with essential links and social icons
- Minimalism
- User-friendly design and easy navigation to main elements.

Submission:

- Use Figma to design the homepage.
- Share the Figma link for review.

Task 3:

Create a creative reel from the post or highlights of DJS Compute and DJS S4DS insta handle

Instagram link:

DJS Compute: https://www.instagram.com/djscompute

DJS S4DS: https://www.instagram.com/djs_s4ds

Feel free to contact any of the mentors below for any queries:

- Charmi: +91 9004992072
- Vidhi: +91 8928769453
- Meet: +91 7021417537

See you at the interviews!