Interview Q&A Guide: ESG GenAI Data Intern Role

Q: What is the difference between a list and a tuple in Python?

A: Lists are mutable (can be changed), while tuples are immutable (cannot be changed).

Q: How do you handle missing data in a dataset using Python?

A: Using pandas functions like dropna(), fillna(), or interpolation methods.

Q: What libraries do you use for data analysis?

A: Common libraries are pandas, NumPy, matplotlib, seaborn, and scikit-learn.

Q: How would you read and process a CSV file in Python?

A: Use pandas: df = pd.read_csv('file.csv'), then clean or transform as needed.

Q: Explain the difference between .loc[] and .iloc[] in pandas.

A: .loc[] accesses rows by labels; .iloc[] accesses rows by index position.

Q: How can you filter rows in a DataFrame based on a condition?

A: Use a boolean condition, e.g., df[df['column'] > 100].

Q: How would you merge or join two DataFrames?

A: Use pd.merge() for relational joins, or pd.concat() to stack them.

Q: What is data validation? How would you approach validating data in Python?

A: Data validation means checking data correctness. Use checks like data types, ranges, or regex for formats.

Q: How do you handle JSON data in Python?

A: Use the json module to load or dump JSON, or pandas.read_json() for tabular formats.

Q: How do you write a simple Python script to find duplicate values in a list or dataset?

A: Use collections.Counter or pandas.duplicated() to identify duplicates.

Q: What is prompt engineering?

A: Its the process of designing input prompts to get better, targeted responses from language models.

Interview Q&A Guide: ESG GenAI Data Intern Role

Q: How does prompt engineering affect the output of an LLM like ChatGPT?

A: It changes clarity, structure, and relevance of the response. Good prompts give better results.

Q: Have you ever tried refining prompts to get better results? Can you give an example?

A: Yes, for summarization tasks I changed 'Summarize this:' to 'Give a 3-bullet summary of:' to improve clarity.

Q: What is the difference between zero-shot, one-shot, and few-shot prompting?

A: Zero-shot: no examples; one-shot: 1 example; few-shot: multiple examples to guide the model.

Q: How would you prompt an LLM to extract structured information from an unstructured paragraph?

A: Use a clear format like: 'Extract the following fields: Name, Age, Location from this text:'

Q: What are some challenges you've faced while using AI tools like ChatGPT or OpenAI API?

A: Challenges include hallucination, inconsistent outputs, and sensitivity to prompt phrasing.

Q: How can you make a prompt more specific or task-focused?

A: Be explicit about format, tone, length, and include examples or structure.

Q: How do you test if a prompt is working well?

A: Check accuracy, consistency, and repeatability of outputs on varied inputs.

Q: What tools or platforms have you used to experiment with LLMs (e.g., OpenAI Playground, Hugging Face)?

A: I've used ChatGPT, OpenAI Playground, and Hugging Face's transformers and inference API.

Q: Can you give an example of a prompt that failed and how you fixed it?

A: A vague 'Write about climate change' gave generic results; rephrased to 'List 3 key causes of climate change with 1-line explanation each' improved it.