**My Prompt**:  
"I'm working on a Digital Humanities project analyzing al-Jazeera articles. I have two DataFrames:

1. tfidf\_df with columns ['filename-1', 'filename-2', 'similarity']
2. topics\_df with ['file', 'topic']

I need to merge topic information from topics\_df into tfidf\_df by matching 'file' to 'filename-1', but keep all original TF-IDF pairs. Can you explain how to do this properly in Pandas without giving me the full code?"

**AI Response**:  
"For your left join operation, you'll want to:

1. Use df.merge() with how='left' to preserve all TF-IDF pairs
2. Rename columns before merging to avoid conflicts
3. Specify the key columns with on= parameter

The basic structure would be:  
merged\_df = left\_df.merge(right\_df[['key\_column', 'desired\_column']].rename(columns={...}), on='key\_column', how='left')

Remember to verify the merge with merged\_df.shape and check for null values."

**AI Documentation Entry #2**  
**Purpose**: Remove unassigned topics (-1) from merged DataFrame

**My Prompt**:  
"How do I filter out rows where topic\_1 or topic\_2 equals -1 in pandas?"

**AI Response**:  
"Use boolean indexing: df[(df['col1'] != -1) & (df['col2'] != -1)].copy(). Check the new shape with df.shape."

**AI#3**

**Prompt:**"I'm analyzing Al Jazeera articles with two DataFrames:

1. tfidf\_df (columns: filename-1, filename-2, similarity)
2. topics\_df (columns: file, topic)  
   How can I merge topic labels into tfidf\_df by matching file to filename-1, while keeping all original similarity pairs?"

AI Response:  
"To preserve all TF-IDF pairs while adding topic data:

1. Perform a left merge (how='left') on filename-1 and file
2. Rename file to filename-1 in topics\_df before merging
3. Select only needed columns (file and topic) from topics\_df
4. Verify results with .shape and null checks  
   Structure:  
   merged\_df = tfidf\_df.merge(topics\_df[['file','topic']].rename(columns={'file':'filename-1'}), on='filename-1', how='left')"

Prompt:  
"Should I handle filename-2 separately after merging filename-1?"

AI Response:  
"Yes, repeat the process for filename-2:

1. Rename file to filename-2 in topics\_df
2. Left-merge again using the new column name
3. Consider renaming topic columns (topic\_1, topic\_2) for clarity"

Prompt:  
"How do I check if the merge worked correctly?"

AI Response:  
"Verify with:

1. merged\_df.shape (should match original row count)
2. merged\_df.isnull().sum() (check unmatched entries)
3. Sample checks with .head()"