

```
import pandas as pd
```

```
# CONFIGURATION
```

```
INPUT_FILE_NAME = "original_data.csv"
```

```
OUTPUT_FILE_NAME = "random_sample_400.csv"
```

```
SAMPLE_SIZE = 400
```

```
RANDOM_SEED = 42
```

```
# EXECUTION
```

```
print("Starting process...")
```

```
# 1. Load the dataset
```

```
#     If the file is not found, Python will stop here with a default error message.
```

```
df = pd.read_csv(INPUT_FILE_NAME, encoding='utf-8-sig')
```

```
print(f"Data loaded. Total records: {len(df)}")
```

```
# 2. Perform random sampling
```

```
if len(df) < SAMPLE_SIZE:
```

```
    sample_df = df.copy()
```

```
else:
```

```
    sample_df = df.sample(n=SAMPLE_SIZE, random_state=RANDOM_SEED)
```

```
# 3. Add columns for human annotation
```

```
sample_df["Human_Sentiment_Label"] = ""
```

```
sample_df["Annotator_Notes"] = ""
```

```
# 4. Save to CSV
```

```
sample_df.to_csv(OUTPUT_FILE_NAME, encoding='utf-8-sig', index=False)
```

```
print(f"Success! Output file created: {OUTPUT_FILE_NAME}")
```