Summary Tab-Data Cleaning

```
import pandas as pd

file_name = "invoice_Annexure_180796_09042025_1744207214805.xlsx"

df=pd.read_excel(file_name,sheet_name='Summary')

data = {

   "Brand": df.iloc[3, 1],

   "Location": df.iloc[4, 1],

   "City": df.iloc[5, 1],

   "Res-ID": df.iloc[6, 1].replace("Rest. ID - ", "") if pd.notna(df.iloc[6, 1]) else None,

   "Payout Period": df.iloc[10, 2],

   "Payout Settlement Date": df.iloc[11, 2],

   "Total Payout": df.iloc[12, 2],

   "Total Orders": df.iloc[13, 2],

   "Bank UTR": df.iloc[14, 2] if pd.notna(df.iloc[14, 2]) else df.iloc[14, 1],# handle if it's in col 1

   "File_Name":file_name

}

result_df = pd.DataFrame([data])

print(result_df)
```

Note:

Here in the **file_name** we will specify the path for all the files and we will extract the information of all details what required. In the similar way we have to write the code by changing the files and getting the information After that we should merge all the files together

The Python Code

```
dfs = [
   result_df, result_df1, result_df2, result_df3, result_df4, result_df5,
   result_df6, result_df7, result_df8, result_df9, result_df10, result_df11
]
merged_df = pd.concat(dfs, ignore_index=True)
data_frame=pd.DataFrame(merged_df)
data_frame
```

Finally we are writing the output to the excel file

```
data_frame.to_excel('Summary_Tab.xlsx')s
```