import pandas as pd

import matplotlib.pyplot as plt

import os

df1 = pd.read\_csv("dataset1.csv") # Bat landings

df2 = pd.read\_csv("dataset2.csv") # Rat arrivals etc.

os.makedirs("figures", exist\_ok=True)

risk\_reward\_ct = pd.crosstab(df1['risk'], df1['reward'], normalize='index') \* 100

avoid\_success = risk\_reward\_ct.loc[0, 1] if 0 in risk\_reward\_ct.index else 0

risk\_success = risk\_reward\_ct.loc[1, 1] if 1 in risk\_reward\_ct.index else 0

risk\_counts = df1['risk'].value\_counts(normalize=False)

risk\_percents = df1['risk'].value\_counts(normalize=True) \* 100

try:

from scipy.stats import chi2\_contingency

contingency = pd.crosstab(df1['risk'], df1['reward'])

chi2, p, dof, expected = chi2\_contingency(contingency)

chi\_result = f"Chi-square = {chi2:.3f}, p-value = {p:.5f}, dof = {dof}"

except ImportError:

chi\_result = "SciPy not available – Chi-square test skipped."

timing\_summary = df1.groupby('risk')['seconds\_after\_rat\_arrival'].describe()

rat\_arrivals = df2['rat\_arrival\_number'].describe()

corr\_rat\_bat = df2[['rat\_arrival\_number','bat\_landing\_number']].corr().iloc[0,1]

monthly\_summary = df2.groupby('month')[['bat\_landing\_number','rat\_arrival\_number']].mean()

plt.figure()

risk\_counts.plot(kind='bar', color=['skyblue','salmon'])

plt.title("Bat Landing Behaviour (Counts)")

plt.xlabel("0=Avoid, 1=Risk-taking")

plt.ylabel("Count")

plt.savefig("figures/risk\_distribution\_counts.png")

plt.close()

plt.figure()

risk\_percents.plot(kind='bar', color=['lightgreen','orange'])

plt.title("Bat Landing Behaviour (Percent)")

plt.xlabel("0=Avoid, 1=Risk-taking")

plt.ylabel("Percentage")

plt.savefig("figures/risk\_distribution\_percent.png")

plt.close()

plt.figure()

risk\_reward\_ct.plot(kind='bar', stacked=True)

plt.title("Risk vs Reward Outcomes (%)")

plt.xlabel("Risk (0=Avoid, 1=Risk-taking)")

plt.ylabel("Percentage")

plt.legend(["No Reward", "Reward"])

plt.savefig("figures/risk\_vs\_reward\_percent.png")

plt.close()

plt.figure()

df1['seconds\_after\_rat\_arrival'].hist(bins=30, color="purple", edgecolor="black")

plt.title("Landing Time After Rat Arrival")

plt.xlabel("Seconds")

plt.ylabel("Frequency")

plt.savefig("figures/seconds\_after\_rat\_arrival\_hist.png")

plt.close()

with open("results\_summary.txt", "w") as f:

f.write("HIT140 – Assessment 2 (Investigation A)\n")

f.write("=====================================\n\n")

f.write("Dataset1 Analysis:\n")

f.write(f"Risk-taking bats: {risk\_counts.get(1,0)} ({risk\_percents.get(1,0):.1f}%)\n")

f.write(f"Avoiding bats: {risk\_counts.get(0,0)} ({risk\_percents.get(0,0):.1f}%)\n\n")

f.write(f"Success rate (avoidance): {avoid\_success:.1f}%\n")

f.write(f"Success rate (risk-taking): {risk\_success:.1f}%\n\n")

f.write(f"Chi-square test: {chi\_result}\n\n")

f.write("Timing analysis (seconds after rat arrival):\n")

f.write(str(timing\_summary))

f.write("\n\n")

f.write("Dataset2 Context:\n")

f.write(f"Rat arrivals per 30 mins (summary):\n{rat\_arrivals}\n\n")

f.write(f"Correlation (rat arrivals vs bat landings): {corr\_rat\_bat:.3f}\n\n")

f.write("Monthly averages (bat landings & rat arrivals):\n")

f.write(str(monthly\_summary))

f.write("\n")

print("Analysis complete. Figures saved in 'figures/' and summary in 'results\_summary.txt'")