

Patryk Ostrowski, mod. 3, zad. 1

Importy

Komentarze

Zmienne

Struktury danych

1. Najłatwiejsze

 pokaż dane rekordu, którego tempo mierzone za każdym razem było najbardziej stabilne

	Open	15 km Tempo	20 km Czas	20 km Miejsce Open	20 km Tempo	Tempo Stabilność	Czas	Tempo	
4,334	4,910	5.6833	01:52:14	4,439	5.6067	0	01:58:24	5.6127	

```
1 # Import the required dependencies
2 import pandas as pd
3
4 # Assuming df is a list of DataFrames and df is the DataFrame we are working with
5 df = df
6
7 # Find the record with the most stable pace
8 most_stable_record = df.loc[df['Tempo Stabilność'].abs().idxmin()]
9
10 # Declare result variable
11 result = {
12     "type": "dataframe",
13     "value": most_stable_record.to_frame().T # Convert the Series to a DataFrame for consistency
14 }
```

2. Średnie



z krajów, z których bieгло więcej niż 10-ciu uczestników wskaż dwóch - jeden z najbardziej stabilnym tempem w każdym pomiarze i jeden z najmniej stabilnym tempem w każdym pomiarze



	m Miejsce Open	15 km Tempo	20 km Czas	20 km Miejsce Open	20 km Tempo	Tempo Stabilność	Czas
4,293	5,842	4.95	01:52:25	4,481	5.0733	-0.1278	01:58:15
1,346	1,923	4.3	01:37:31	1,428	4.69	-0.0562	01:42:31
2,916	3,292	5.2833	01:46:00	2,976	5.3433	0.0005	01:51:57
5,461	5,945	5.8167	01:57:03	5,560	5.93	-0.0007	02:03:33
5,894	6,539	6.0333	01:59:25	5,999	5.9733	-0.0025	02:05:57
78	98	3.9733	01:18:24	78	3.95	0.0027	01:22:52
8,022	3,746	7.7967	02:12:06	8,156	10.2533	0.4361	02:18:56
9,764	9,540	7.9233	02:30:11	9,720	8.68	0.1739	02:40:41
9,853	9,512	8.01	02:31:08	9,761	8.9167	0.1721	02:42:42
10,055	9,800	7.9267	02:38:25	10,019	9.6567	0.1743	02:50:10



```

# Import the required dependencies
import pandas as pd

# Assuming df is a list of DataFrames and we are working with the first DataFrame
df = df

# Filter out countries with more than 10 participants
country_counts = df['Kraj'].value_counts()
countries_with_more_than_10 = country_counts[country_counts > 10].index

# Initialize lists to store results
most_stable_participants = []
least_stable_participants = []

# Iterate over each country
for country in countries_with_more_than_10:
    # Filter participants from the current country
    country_participants = df[df['Kraj'] == country]

    # Find the participant with the most stable pace
    most_stable = country_participants.loc[country_participants['Tempo Stabilność'].idxmin()]
    most_stable_participants.append(most_stable)

    # Find the participant with the least stable pace
    least_stable = country_participants.loc[country_participants['Tempo Stabilność'].idxmax()]
    least_stable_participants.append(least_stable)

# Combine results into a single DataFrame
result_df = pd.DataFrame(most_stable_participants + least_stable_participants)

```

3. Najtrudniejsze



podaj odchylenie standardowe dla średniej wartości tempa z każdego pomiaru dla uczestników, którzy zajęli dwadzieścia pierwszych miejsc i dla tych z 50-ciu ostatnich miejsc



	Group	Standard Deviation
0	Top 21	0.1218
1	Bottom 50	0.805



```
# TODO: import the required dependencies
import pandas as pd

# Load the dataframe
df = df

# Filter the top 21 and bottom 50 participants based on 'Miejsce'
top_21 = df.nsmallest(21, 'Miejsce')
bottom_50 = df.nlargest(50, 'Miejsce')

# Calculate the mean pace for each measurement point for both groups
top_21_mean_pace = top_21[['5 km Tempo', '10 km Tempo', '15 km Tempo', '20 km Tempo']].mean()
bottom_50_mean_pace = bottom_50[['5 km Tempo', '10 km Tempo', '15 km Tempo', '20 km Tempo']].mean()

# Calculate the standard deviation of the mean pace for each group
top_21_std_dev = top_21_mean_pace.std()
bottom_50_std_dev = bottom_50_mean_pace.std()

# Declare result var
result = {
    "type": "dataframe",
    "value": pd.DataFrame({
        "Group": ["Top 21", "Bottom 50"],
        "Standard Deviation": [top_21_std_dev, bottom_50_std_dev]
    })
}
```

Co ciekawe w tym zadaniu użyłem liczebnika głównego ‘dwadzieścia’ i zbiorowego ‘pierwszych’ i model odczytał to jako ‘dwadzieścia jeden’, a następnie oparł swoje obliczenia na założeniu 21 rekordów. Zostawiłem tak, bo i tak jest to konsekwentne w kodzie. Jednak w celu wyłuskania korzyści biznesowej dla zamawiającego, należałoby to doprecyzować – to też element uczenia się języka konkretnego modelu i dowód na konieczność weryfikowania zwróconych wyników.