IU South Bend Student Enrollment Appendix

March 30, 2024

[1]:
```
by Pedro H.
```

[ ]:
```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
from scipy import stats
import seaborn as sns

file_path = r"C:\Users\pedro\Downloads\Untitled Folder\IU South Bend Beginner␣
 ↪Enrollment by Race and Ethnicity 2012 to 2023.xlsx"

iu_new_s_data = pd.read_excel(file_path)
```

[2]:
```
iu_new_s_data = iu_new_s_data.rename(columns={"Unnamed: 0": "Year"})
iu_new_s_data = iu_new_s_data.T
iu_new_s_data.reset_index(inplace=True)
iu_new_s_data = iu_new_s_data.rename(columns=iu_new_s_data.iloc[0]).drop(0)
iu_new_s_data["Domestic"] = iu_new_s_data["Total, Domestic Known"] +␣
 ↪iu_new_s_data["Domestic Unknown"]
iu_new_s_data = iu_new_s_data.dropna(axis=1)
iu_new_s_data = iu_new_s_data.astype(int)
print(iu_new_s_data.head(13))
```

```
    Year  Hispanic/Latino  African American  Two or More Races  \
1   2012               85                71                 52
2   2013               81                42                 35
3   2014              110                65                 32
4   2015               95                56                 48
5   2016              108                73                 37
6   2017              139                90                 61
7   2018              142                74                 63
8   2019              185                54                 55
9   2020              161                64                 57
10  2021              151                56                 31
11  2022              207                65                 48
12  2023              240                75                 46

    Underrepresented Total  Asian American  Students of Color Total  White  \
```

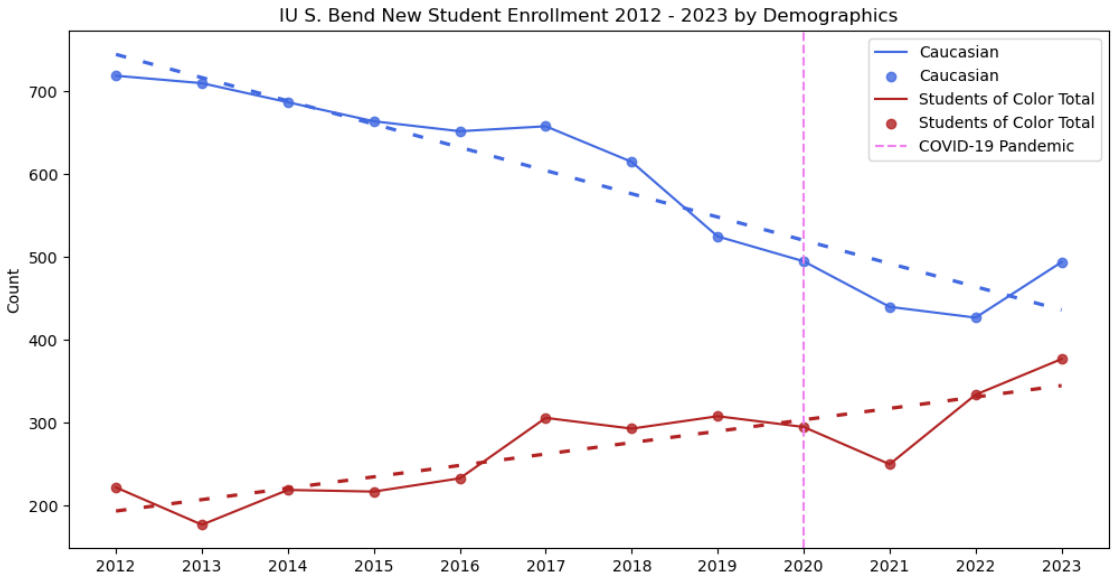
1	211	11	222	719
2	161	16	177	710
3	209	10	219	687
4	202	15	217	664
5	222	11	233	652
6	293	13	306	658
7	280	13	293	615
8	296	12	308	525
9	283	12	295	495
10	239	11	250	440
11	322	12	334	427
12	366	11	377	494

	Total, Domestic Known	Domestic Unknown	International	\
1	941	9	5	
2	887	9	25	
3	906	7	31	
4	881	2	23	
5	885	6	22	
6	964	6	12	
7	908	5	16	
8	833	6	12	
9	790	5	10	
10	690	1	6	
11	761	2	8	
12	871	3	12	

	Total Beginner Students	Domestic
1	955	950
2	921	896
3	944	913
4	906	883
5	913	891
6	982	970
7	929	913
8	851	839
9	805	795
10	697	691
11	771	763
12	886	874

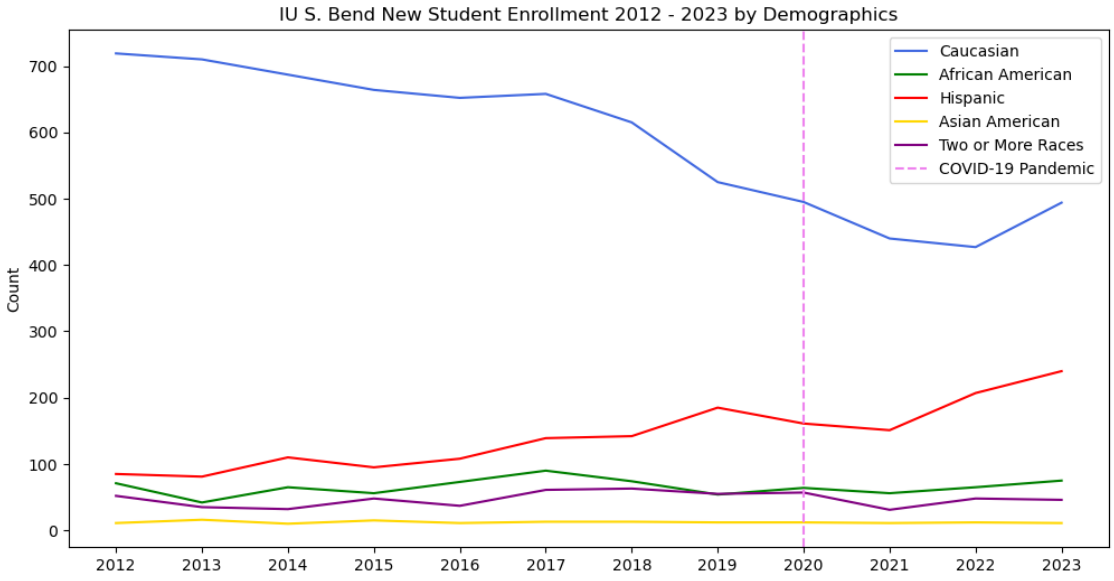
[3]:
```
plt.figure(figsize=(12, 6))

sns.lineplot(data=iu_new_s_data, x='Year', y='White', label='Caucasian',␣
 ↪color='royalblue')
sns.regplot(data=iu_new_s_data, x='Year', y='White', label='Caucasian',␣
 ↪line_kws={'dashes': (3, 5)}, ci=None, color='royalblue')
```

```
sns.lineplot(data=iu_new_s_data, x='Year', y='Students of Color Total',␣
 ↪label='Students of Color Total', color='firebrick')
sns.regplot(data=iu_new_s_data, x='Year', y='Students of Color Total',␣
 ↪label='Students of Color Total', line_kws={'dashes': (3, 5)}, ci=None,␣
 ↪color='firebrick')

plt.axvline(x=2020, color='violet', linestyle='--', label='COVID-19 Pandemic')

plt.ylabel('Count')
plt.title('IU S. Bend New Student Enrollment 2012 - 2023 by Demographics')
plt.xticks(iu_new_s_data['Year'])
plt.xlabel('')

plt.legend()
plt.show()
```



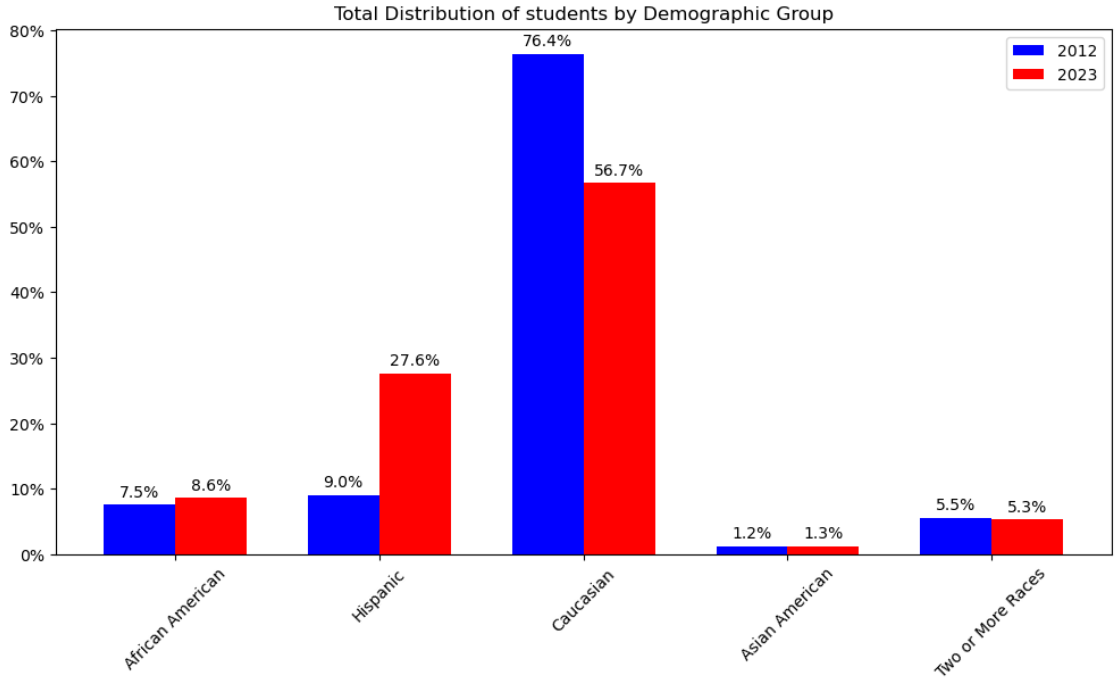
[4]:
```
plt.figure(figsize=(12, 6))

sns.lineplot(data=iu_new_s_data, x='Year', y='White', label='Caucasian',␣
 ↪color='royalblue')
sns.lineplot(data=iu_new_s_data, x='Year', y='African American', label='African␣
 ↪American', color='green')
sns.lineplot(data=iu_new_s_data, x='Year', y='Hispanic/Latino',␣
 ↪label='Hispanic', color='red')
sns.lineplot(data=iu_new_s_data, x='Year', y='Asian American', label='Asian␣
 ↪American', color='gold')
```

```
sns.lineplot(data=iu_new_s_data, x='Year', y='Two or More Races', label='Two or␣
 ↪More Races', color='purple')

plt.axvline(x=2020, color='violet', linestyle='--', label='COVID-19 Pandemic')

plt.ylabel('Number of students')
plt.title('IU S. Bend New Student Enrollment 2012 - 2023 by Demographics␣
 ↪(Disaggregated)')
plt.xticks(iu_new_s_data['Year'])
plt.xlabel('')

plt.legend()
plt.show()
```



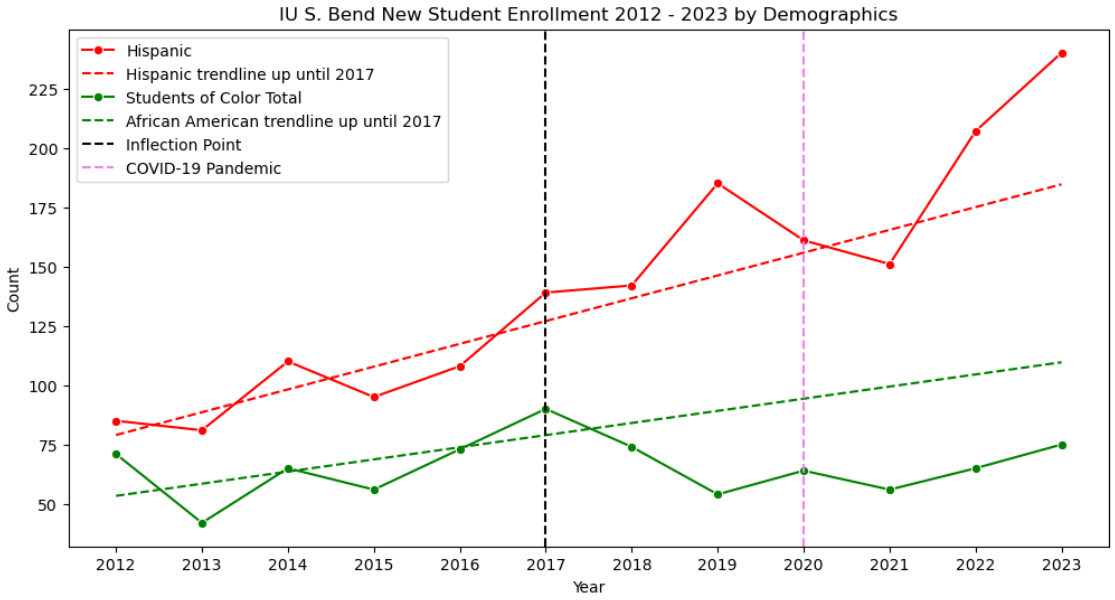
[14]:
```
ethnicities = ['African American', 'Hispanic', 'Caucasian', 'Asian American',␣
 ↪'Two or More Races']

iu_new_s_data_2012 = iu_new_s_data[iu_new_s_data['Year'] == 2012]
iu_new_s_data_2023 = iu_new_s_data[iu_new_s_data['Year'] == 2023]

total_known_2012 = iu_new_s_data_2012['Total, Domestic Known'].sum()
values_2012 = [
    (iu_new_s_data_2012['African American'].sum() / total_known_2012) * 100,
    (iu_new_s_data_2012['Hispanic/Latino'].sum() / total_known_2012) * 100,
    (iu_new_s_data_2012['White'].sum() / total_known_2012) * 100,
    (iu_new_s_data_2012['Asian American'].sum() / total_known_2012) * 100,
    (iu_new_s_data_2012['Two or More Races'].sum() / total_known_2012) * 100
```

```
]

total_known_2023 = iu_new_s_data_2023['Total, Domestic Known'].sum()
values_2023 = [
    (iu_new_s_data_2023['African American'].sum() / total_known_2023) * 100,
    (iu_new_s_data_2023['Hispanic/Latino'].sum() / total_known_2023) * 100,
    (iu_new_s_data_2023['White'].sum() / total_known_2023) * 100,
    (iu_new_s_data_2023['Asian American'].sum() / total_known_2023) * 100,
    (iu_new_s_data_2023['Two or More Races'].sum() / total_known_2023) * 100
]

plt.figure(figsize=(12, 6))
bar_width = 0.35
index = np.arange(len(ethnicities))

bar1 = plt.bar(index, values_2012, bar_width, label='2012', color='b')
bar2 = plt.bar(index + bar_width, values_2023, bar_width, label='2023',␣
 ↪color='r')

plt.xlabel('')
plt.ylabel('')
plt.title('Total Distribution of students by Demographic Group')
plt.xticks(index + bar_width / 2, ethnicities, rotation=45)
plt.legend()

plt.yticks(np.arange(0, max(max(values_2012), max(values_2023)) + 10, 10))
plt.gca().set_yticklabels(['{:.0f}%'.format(x) for x in plt.gca().get_yticks()])

for bars in [bar1, bar2]:
    for bar in bars:
        height = bar.get_height()
        plt.annotate('{:.1f}%'.format(height),
                     xy=(bar.get_x() + bar.get_width() / 2, height),
                     xytext=(0, 3),
                     textcoords="offset points",
                     ha='center', va='bottom')

plt.show()
```



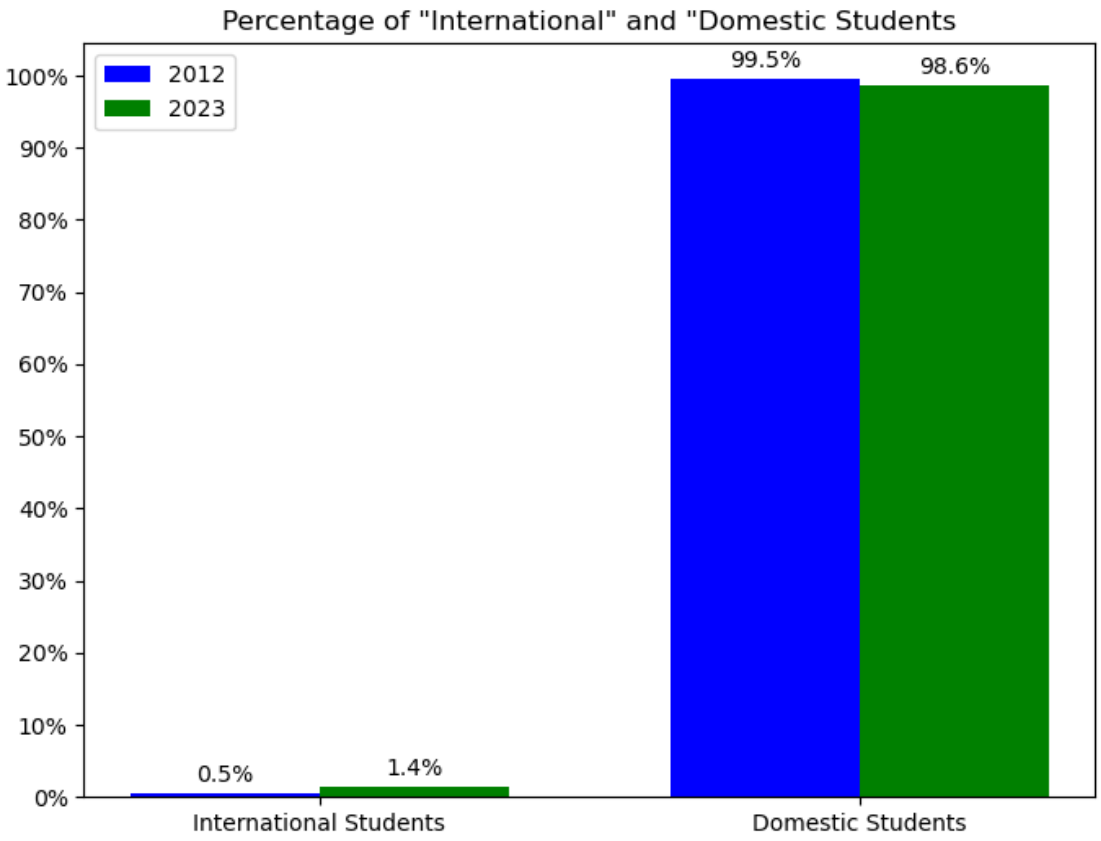
[6]:
```
iu_new_s_data_filtered = iu_new_s_data[iu_new_s_data['Year'] <= 2017]

plt.figure(figsize=(12, 6))

sns.lineplot(data=iu_new_s_data, x='Year', y='Hispanic/Latino',␣
 ↪label='Hispanic', color='red', marker='o')

slope_hispanic, intercept_hispanic, _, _, _ = stats.
 ↪linregress(iu_new_s_data_filtered['Year'], iu_new_s_data_filtered['Hispanic/
 ↪Latino'])
x_values_hispanic = np.array([min(iu_new_s_data['Year']), 2023])
y_values_hispanic = slope_hispanic * x_values_hispanic + intercept_hispanic
plt.plot(x_values_hispanic, y_values_hispanic, '--', color='red',␣
 ↪label=f'Hispanic trendline up until 2017')

sns.lineplot(data=iu_new_s_data, x='Year', y='African American',␣
 ↪label='Students of Color Total', color='green', marker='o')

slope_african_american, intercept_african_american, _, _, _ = stats.
 ↪linregress(iu_new_s_data_filtered['Year'], iu_new_s_data_filtered['African␣
 ↪American'])
x_values_african_american = np.array([min(iu_new_s_data['Year']), 2023])
y_values_african_american = slope_african_american * x_values_african_american␣
 ↪+ intercept_african_american
```

```
plt.plot(x_values_african_american, y_values_african_american, '--',␣
 ↪color='green', label=f'African American trendline up until 2017')

plt.axvline(x=2017, color='black', linestyle='--', label='Inflection Point')
plt.axvline(x=2020, color='violet', linestyle='--', label='COVID-19 Pandemic')

plt.ylabel('Number of Students')
plt.title('IU S. Bend New Student Enrollment 2012 - 2023 by Demographics')
plt.xlabel('')
plt.xticks(iu_new_s_data[''])

plt.legend()
plt.show()
```



[15]:
```
iu_new_s_data_2012 = iu_new_s_data[iu_new_s_data['Year'] == 2012]
iu_new_s_data_2023 = iu_new_s_data[iu_new_s_data['Year'] == 2023]

total_known_2012 = iu_new_s_data_2012['Total Beginner Students'].sum()
values_2012 = [
    (iu_new_s_data_2012['International'].sum() / total_known_2012) * 100,
    (iu_new_s_data_2012['Domestic'].sum() / total_known_2012) * 100
]

total_known_2023 = iu_new_s_data_2023['Total Beginner Students'].sum()
values_2023 = [
    (iu_new_s_data_2023['International'].sum() / total_known_2023) * 100,
    (iu_new_s_data_2023['Domestic'].sum() / total_known_2023) * 100
```

```
]

plt.figure(figsize=(8, 6))
bar_width = 0.35
index = np.arange(2)

bar1 = plt.bar(index, values_2012, bar_width, label='2012', color='b')
bar2 = plt.bar(index + bar_width, values_2023, bar_width, label='2023',␣
 ↪color='g')

plt.xlabel('')
plt.ylabel('')
plt.title('Percentage of "International" and "Domestic Students')
plt.xticks(index + bar_width / 2, ['International Students', 'Domestic␣
 ↪Students'])
plt.legend()

plt.yticks(np.arange(0, max(max(values_2012), max(values_2023)) + 10, 10))
plt.gca().set_yticklabels(['{:.0f}%'.format(x) for x in plt.gca().get_yticks()])

for bars in [bar1, bar2]:
    for bar in bars:
        height = bar.get_height()
        plt.annotate('{:.1f}%'.format(height),
                     xy=(bar.get_x() + bar.get_width() / 2, height),
                     xytext=(0, 3),
                     textcoords="offset points",
                     ha='center', va='bottom')

plt.show()
```



[ ]:
```
plt.figure(figsize=(12, 6))

sns.lineplot(data=iu_new_s_data, x='Year', y='International',␣
 ↪label='International', color='forestgreen')
sns.regplot(data=iu_new_s_data, x='Year', y='International',␣
 ↪label='International', line_kws={'dashes': (3, 5)}, ci=None,␣
 ↪color='forestgreen')

plt.axvline(x=2020, color='violet', linestyle='--', label='COVID-19 Pandemic')

plt.ylabel('Count')
plt.title('IU S. Bend New Student Enrollment 2012 - 2023 by Demographics')
plt.xticks(iu_new_s_data['Year'])
plt.xlabel('')

plt.legend()
plt.show()
```

[ ]:
```

```