

Understanding catplot Function

```
# Step 1: Melt the DataFrame to long format
df_melted = df.melt(
    id_vars=["country_name", "continent"],
    value_vars=[str(year) for year in range(2010, 2022)],
    var_name="year",
    value_name="unemployment_rate"
)

# Step 2: Calculate median unemployment rate by
continent and year
median_df = df_melted.groupby(['continent', 'year'])
['unemployment_rate'].median().reset_index()

# Step 3: Plot
plt.figure(figsize=(12, 6))
sns.lineplot(data=median_df, x="year",
y="unemployment_rate", hue="continent", marker="o")

# Customize the plot
plt.title("Median Unemployment Rate by Continent (2010-
2021)", fontsize=14)
plt.xlabel("Year")
plt.ylabel("Median Unemployment Rate (%)")
plt.xticks(rotation=45)
plt.grid(True)
plt.legend(title="Continent", bbox_to_anchor=(1.05, 1),
loc='upper left')
plt.tight_layout()
plt.show()
```

why need to reset index

Great question!

In your code:

python

```
median_df = df_melted.groupby(['continent', 'year'])
['unemployment_rate'].median().reset_index()
```

the key part is `reset_index()` — and here's why you need it:

🔍 What happens without `reset_index()` ?

When you run:

```
python
```

```
df_melted.groupby(['continent', 'year'])['unemployment_rate'].median()
```

You get a **Series with a MultiIndex** — `continent` and `year` are not columns anymore; they're part of the index.

That looks like this:

```
yaml
```

```
continent    year
Africa      2010    5.1
           2011    5.3
Asia        2010    4.2
           2011    4.0
Name: unemployment_rate, dtype: float64
```

That kind of object is **not convenient** for plotting with seaborn or pandas — seaborn expects columns, not MultiIndexes.

✅ What does `reset_index()` do?

It converts the MultiIndex into regular columns:

```
python
```

```
continent  year  unemployment_rate
0    Africa  2010             5.1
1    Africa  2011             5.3
2     Asia  2010             4.2
3     Asia  2011             4.0
```

Now you have a **clean, flat DataFrame** with proper columns, which `sns.lineplot()` can easily use.

Summary

- `reset_index()` is used to **move index levels** (like 'continent' and 'year') **back into columns**.
- This makes the data easier to **view, manipulate, and plot**, especially with seaborn or matplotlib.

Would you like to see what happens visually if you omit `reset_index()` in this context?