**What is a Stationarity Test — and Why It’s Crucial for Predictive Modeling?**

A **stationarity test** is a statistical procedure used to determine whether a **time series** has **constant statistical properties** (like **mean**, **variance**, and **autocorrelation**) over time.

**Stationarity** refers to a time series whose **mean, variance, and autocorrelation remain constant over time**. It's a fundamental assumption for many predictive models like **ARIMA** and **SARIMA**.

If these properties change over time (non-stationary), forecasting becomes unreliable unless the data is transformed.

**❓ Why Do We Perform a Stationarity Test?**

Because most **time series forecasting models (like ARIMA, SARIMA, etc.) assume stationarity** in the input data.

**Key Reasons:**

* 🔄 **Ensure model assumptions are met**.
* 📉 **Improve forecasting accuracy** by stabilizing variance and trend.
* 🔍 **Identify need for transformation** (e.g., differencing, detrending, or logging).

**🔍 Why is this important?**

➡ Stationary data ensures model stability and more reliable long-term forecasts.  
➡ If the data is non-stationary, we apply **transformations** (like differencing) to stabilize it.  
➡ This test helps determine the **right modeling strategy**—saving time and improving accuracy.

**⏱️ When Do We Perform a Stationarity Test?**

We perform it **before fitting a predictive time series model**, typically after:

* Performing **initial EDA (exploratory data analysis)**.
* Plotting the **time series** and observing patterns.
* Noticing **trends, seasonality**, or **drifting variance**.

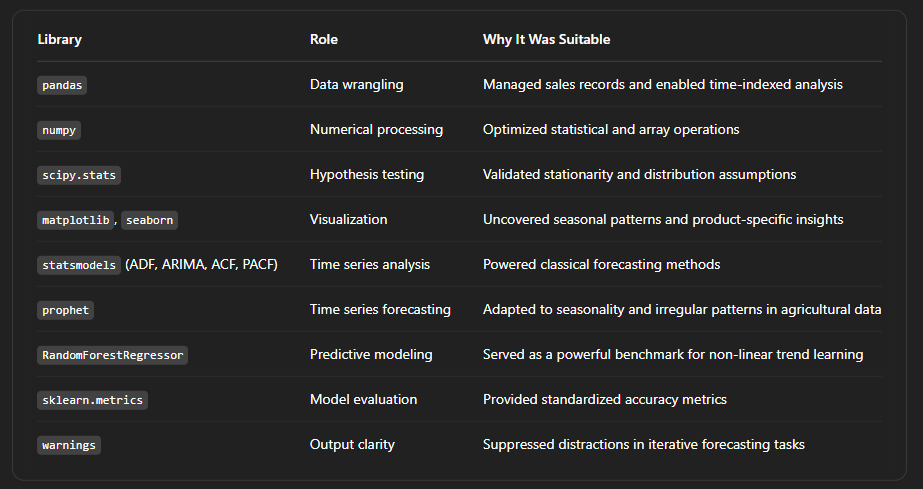
**🧠 What Role Does Stationarity Play in Predictive Modeling?**

In models like ARIMA:

* A **stationary time series** simplifies modeling.
* Non-stationary series may lead to **biased coefficients**, **inaccurate confidence intervals**, and **unreliable forecasts**.

A stationary series ensures:

* Better **parameter estimation**.
* More **reliable predictions**.
* Easier detection of **true underlying signals** in the data.



Interpreting the ADF Test Code

# Key part of result:

# p-value < 0.05 → Reject H0 → Data is stationary

# p-value ≥ 0.05 → Fail to reject H0 → Data is non-stationary

This means:

* If the **p-value is low**, your time series **does not have a unit root**, hence it is **stationary**.
* If the **p-value is high**, it **needs transformation** before modeling (like differencing or removing trend).

**🛠️ Technologies & Tools Used:**

* **Python**
* **Pandas**, **Statsmodels**
* **ADF Test** for unit root detection
* **Visualization** for trend inspection
* **Train/Test trend analysis** to assess model performance