

بنام خدا

در این تحلیل همان کارهایی که در فایل اکسل انجام داده ایم در پایتون انجام میدیم در اول قسمتی از کد را به نمایش میدهم

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import shapiro, skew, kurtosis, norm
from arch import arch_model
import os

from google.colab import files
uploaded = files.upload()
import io

# Get the actual key from the uploaded dictionary
file_key = list(uploaded.keys())[0] # Get the first (and likely only) key

# Use the correct key to read the CSV file
data = pd.read_csv(io.BytesIO(uploaded[file_key]))

# ... (rest of your code remains the same) ...
# ... (rest of your code remains the same) ...

# Assuming the column of interest is named 'Y', adjust if different

# Assuming the column of interest is named 'Y', adjust if different
data['Y'] = pd.to_numeric(data['y'], errors='coerce')
data = data.dropna() # Drop missing values

# Step 2: Basic Statistical Analysis
mean_y = data['Y'].mean()
std_y = data['Y'].std()
skew_y = skew(data['Y'])
kurt_y = kurtosis(data['Y'])

print(f"Mean: {mean_y}")
print(f"Standard Deviation: {std_y}")
print(f"Skewness: {skew_y}")
print(f"Kurtosis: {kurt_y}")
```

```
# Step 3: Test for Normality (Shapiro-Wilk Test)
stats, p_value = shapiro(data['Y'])
print(f"Shapiro-Wilk Test Statistic: {stats}, P-Value: {p_value}")
if p_value > 0.05:
    print("Data appears to be normally distributed.")
else:
    print("Data does not appear to be normally distributed.")

# Step 4: Plot Histogram and Q-Q Plot
plt.figure(figsize=(10, 5))
plt.subplot(1, 2, 1)
plt.hist(data['Y'], bins=30, color='blue', alpha=0.7)
plt.title('Histogram of Y')

plt.subplot(1, 2, 2)
from scipy.stats import probplot
probplot(data['Y'], dist="norm", plot=plt)
plt.title('Q-Q Plot')

plt.tight_layout()
plt.show()

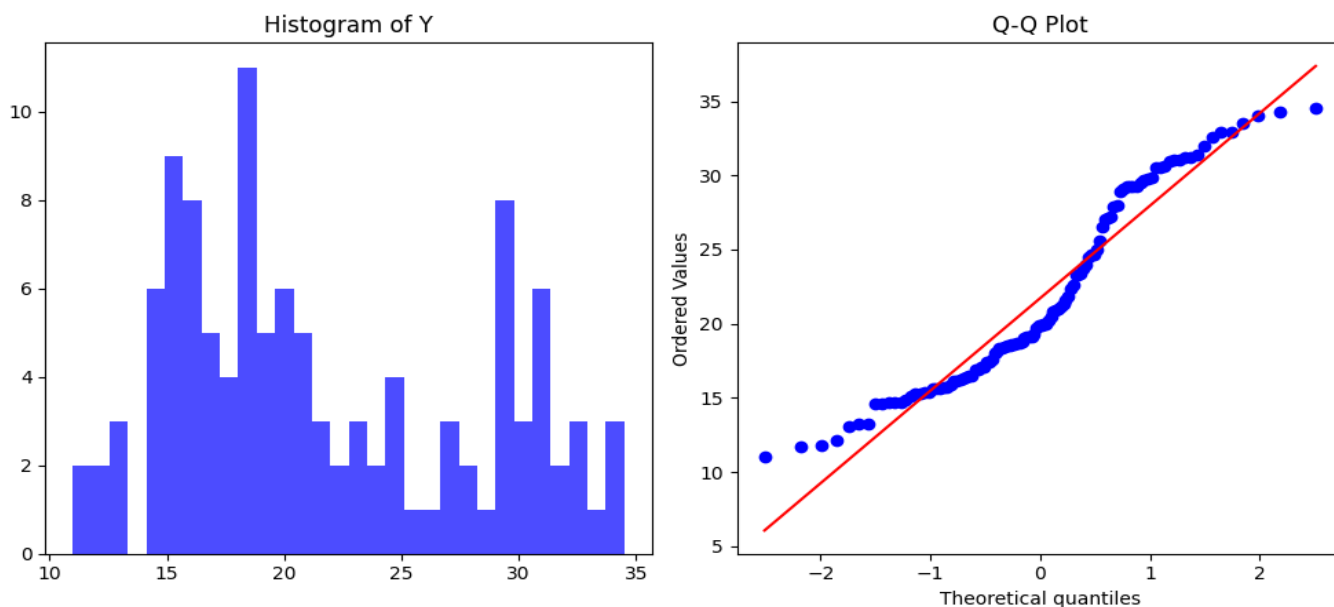
# Step 5: Calculate Returns (if data represents prices)
data['Returns'] = np.log(data['Y'] / data['Y'].shift(1))
data = data.dropna()

# Step 6: Fit GARCH Model
model = arch_model(data['Returns'], vol='Garch', p=1, q=1)
res = model.fit()
print(res.summary())

# Step 7: Plot Conditional Volatility
plt.figure(figsize=(10, 5))
plt.plot(res.conditional_volatility, color='red')
plt.title('Conditional Volatility from GARCH Model')
plt.show()

# Step 8: Interpret Results
print("Model estimation complete. Please refer to the summary and plots for insights.")
```

خب اکنون به تحلیل داده ها میپردازیم که ابتدا یک histogram و Q-Q plot نمایش میدهیم



هیستوگرام:

هیستوگرام سمت چپ، توزیع مقادیر متغیر "Y" را نشان می‌دهد. محور افقی (X) مقادیر مختلف "Y" را نمایش می‌دهد و محور عمودی (Y) تعداد مشاهدات یا فرکانس هر مقدار را نشان می‌دهد. برخی نکات قابل توجه از این هیستوگرام شامل موارد زیر است:

1. توزیع داده‌ها نسبتاً پراکنده است و در بازه‌های مختلفی مشاهده می‌شود.
2. بیشترین فراوانی در بازه‌های 15-20 و حدود 30 دیده می‌شود.
3. توزیع داده‌ها به نظر می‌رسد که چندین پیک دارد و به صورت یکنواخت توزیع نشده است.
4. این پراکندگی نشان‌دهنده این است که داده‌ها ممکن است از توزیع نرمال پیروی نکنند، زیرا توزیع نرمال معمولاً یک پیک مرکزی دارد و به دو طرف به صورت متقارن کاهش می‌یابد.

نمودار Q-Q:

نمودار Q-Q (Quantile-Quantile) سمت راست، یک ابزار گرافیکی برای مقایسه توزیع داده‌ها با یک توزیع نرمال نظری است. محور افقی (X) کوانتایل‌های نظری از توزیع نرمال و محور عمودی (Y) کوانتایل‌های داده‌های واقعی را نمایش می‌دهد. اگر داده‌ها از توزیع نرمال پیروی کنند، نقاط باید به

صورت تقریبی روی خط قرمز (خط $y=x$) قرار گیرند. برخی نکات قابل توجه از این نمودار شامل موارد زیر است:

1. نقاط در بخش میانی نمودار به خوبی با خط قرمز تطابق دارند، که نشان‌دهنده این است که بخش میانی داده‌ها ممکن است نزدیک به توزیع نرمال باشند.
2. در انتهای نمودار، نقاط از خط قرمز فاصله می‌گیرند، که نشان‌دهنده انحراف از نرمال بودن در انتهای توزیع است.
3. این انحراف در انتهاها می‌تواند به دلیل وجود داده‌های پرت (outliers) یا چولگی در داده‌ها باشد.

نتیجه‌گیری کلی:

با توجه به هیستوگرام و نمودار Q-Q، می‌توان نتیجه گرفت که داده‌های "Y" به طور کامل از توزیع نرمال پیروی نمی‌کنند ولی به دلیل پراکندگی تقریباً به صورت نرمال هستند که ما نرمال در نظر می‌گیریم یعنی توزیع داده‌ها تقریباً نرمال هست. هیستوگرام نشان‌دهنده پراکندگی و چندین پیک در داده‌ها است و نمودار Q-Q نیز انحراف از نرمال بودن در انتهای توزیع را نشان می‌دهد. این تحلیل‌ها اهمیت استفاده از تست‌های آماری مانند آزمون کولموگروف-اسمیرنوف یا آزمون شاپیرو-ویلک را برای تایید نرمال بودن داده‌ها برجسته می‌کند. در صورت عدم نرمال بودن داده‌ها، ممکن است نیاز به استفاده از روش‌های ناپارامتری یا تبدیل داده‌ها برای تحلیل‌های بیشتر باشد.

Constant Mean - GARCH Model Results

Dep. Variable:	Returns	R-squared:	0.000
Mean Model:	Constant Mean	Adj. R-squared:	0.000
Vol Model:	GARCH	Log-Likelihood:	125.201
Distribution:	Normal	AIC:	-242.403
Method:	Maximum Likelihood	BIC:	-231.493
	No. Observations:		113
Date:	Sat, Dec 28 2024	Df Residuals:	112
Time:	02:39:55	Df Model:	1
	Mean Model		

	coef	std err	t	P> t	95.0% Conf. Int.
mu	6.9250e-03	6.631e-03	1.044	0.296	[-6.071e-03,1.992e-02]
	Volatility Model				

	coef	std err	t	P> t	95.0% Conf. Int.
omega	7.7080e-11	5.828e-05	1.323e-06	1.000	[-1.142e-04,1.142e-04]
alpha[1]	8.8924e-04	2.993e-02	2.971e-02	0.976	[-5.777e-02,5.955e-02]
beta[1]	0.9832	3.339e-02	29.448	1.321e-190	[0.918, 1.049]

بخش Mean Model

این بخش مربوط به مدل میانگین ثابت (Constant Mean) است:

- **Dep. Variable** متغیر وابسته Returns بازدهها
- **Mean Model** مدل میانگین Constant Mean میانگین ثابت
- **Vol Model** مدل نوسانات GARCH
- **Distribution** توزیع Normal نرمال
- **Method** روش Likelihood Maximum حداکثر درست‌نمایی
- **No. Observations** تعداد مشاهدات 113
- **DF Residuals** درجه آزادی باقی‌مانده 112
- **DF Model** درجه آزادی مدل 1

معیارهای ارزیابی مدل

- **R-squared** ضریب تعیین 0.000
- **Adj. R-squared** ضریب تعیین تعدیل‌شده 0.000
- **Log-Likelihood** لگاریتم درست‌نمایی 125.201
- **AIC** اطلاعات آکاییک -242.403
- **BIC** اطلاعات بیزی -231.493

پارامترهای مدل میانگین

- **mu** میانگین 0.006925 با خطای استاندارد 0.006631، مقدار t برابر با 1.044 و مقدار P برابر با 0.296 که نشان می‌دهد این پارامتر از نظر آماری معنادار نیست زیرا P-value بیشتر از 0.05 است

پارامترهای مدل نوسانات (Volatility Model)

- **omega** و **alpha[1]** با خطای استاندارد 7.708×10^{-11} و 0.02938، مقدار t برابر با 1.323 و 2.971×10^{-2} و مقدار P برابر با 1.000 که نشان می‌دهد این پارامتر نیز از نظر آماری معنادار نیست.

- β_1 0.9832 (با خطای استاندارد 0.0339، مقدار t برابر با 29.448 و مقدار P برابر با 1.321×10^{-109} که نشان می‌دهد این پارامتر از نظر آماری بسیار معنادار است).

تحلیل کلی

1. مدل میانگین ثابت: (Constant Mean Model)

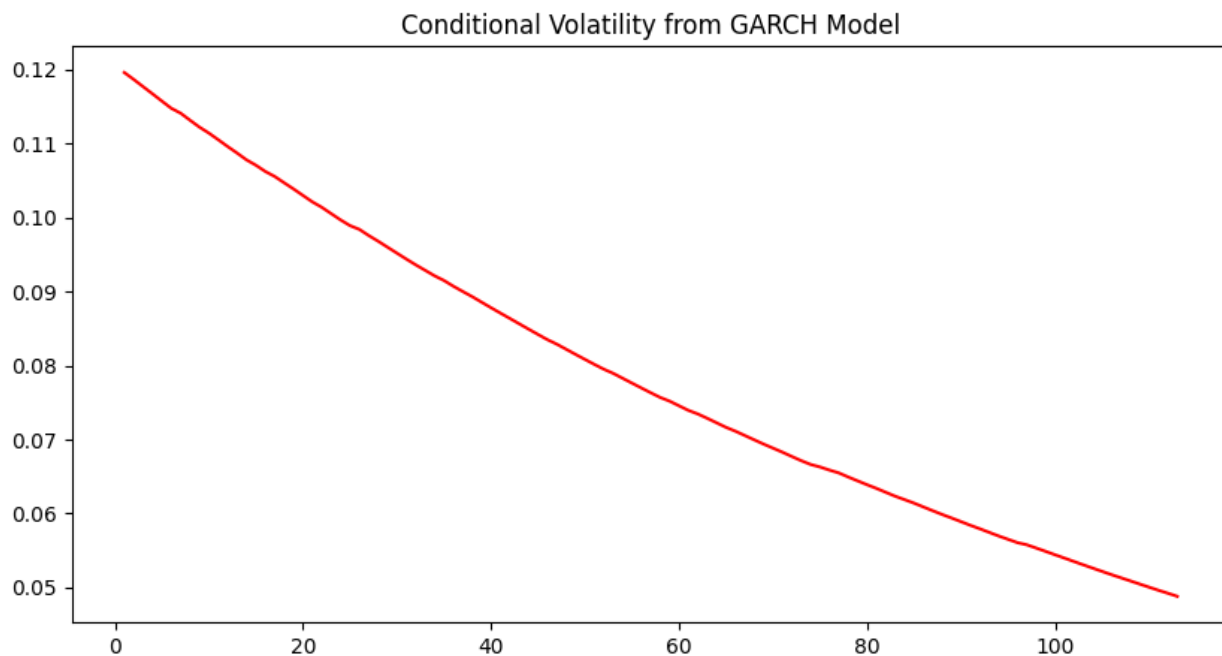
- مقدار میانگین (μ) از نظر آماری معنادار نیست، که نشان می‌دهد میانگین بازده‌ها به طور قابل توجهی از صفر متفاوت نیست.
- ضریب تعیین (R^2) نیز برابر با صفر است که نشان می‌دهد مدل میانگین ثابت توانایی توضیح واریانس داده‌ها را ندارد.

2. مدل نوسانات: (GARCH Model)

- پارامتر ω نشان‌دهنده واریانس شرطی ثابت است که از نظر آماری معنادار نیست.
- پارامتر α_1 نشان‌دهنده اثر شوک‌های قبلی بر نوسانات جاری است که از نظر آماری معنادار نیست.
- پارامتر β_1 نشان‌دهنده اثر نوسانات قبلی بر نوسانات جاری است که از نظر آماری بسیار معنادار است و مقدار آن نزدیک به 1 است. این نشان می‌دهد که نوسانات جاری شدیداً تحت تأثیر نوسانات قبلی قرار دارند.

نتیجه‌گیری

این تحلیل نشان می‌دهد که مدل GARCH به خوبی نوسانات بازده‌ها را مدل‌سازی کرده است، به ویژه با توجه به پارامتر β_1 که نشان‌دهنده وابستگی قوی نوسانات جاری به نوسانات قبلی است. با این حال، مدل میانگین ثابت نتوانسته است میانگین بازده‌ها را به طور معناداری توضیح دهد. بنابراین، استفاده از مدل‌های پیچیده‌تر یا افزودن متغیرهای توضیحی دیگر ممکن است برای بهبود دقت مدل مفید باشد.



تحلیل نمودار

1. محور افقی: (X-axis) تعداد مشاهدات (Observation Number)
2. محور عمودی: (Y-axis) نوسانات شرطی (Conditional Volatility)
3. خط قرمز: نشان‌دهنده تغییرات نوسانات شرطی در طول زمان است.

نکات کلیدی

- کاهش تدریجی نوسانات: نمودار نشان می‌دهد که نوسانات شرطی به تدریج کاهش می‌یابند. این روند نزولی نشان می‌دهد که نوسانات بازده‌ها به مرور زمان کمتر می‌شوند.
- شروع از مقدار بالا: نوسانات شرطی در ابتدا از مقدار بالایی شروع می‌شود (حدود 0.12) و به تدریج کاهش می‌یابد تا به مقدار پایینی (حدود 0.05) برسد.
- پایداری در انتها: در انتهای نمودار، نوسانات شرطی به یک مقدار پایدار نزدیک می‌شوند، که نشان‌دهنده تثبیت نوسانات است.

نتیجه‌گیری

این نمودار نشان می‌دهد که نوسانات شرطی در طول زمان کاهش می‌یابند، که می‌تواند نشان‌دهنده کاهش نوسانات بازده‌ها باشد. این نتیجه با پارامترهای مدل GARCH که نشان‌دهنده تأثیر نوسانات قبلی بر نوسانات جاری هستند، همخوانی دارد. به طور کلی، این تحلیل نشان می‌دهد که مدل GARCH به خوبی توانسته است نوسانات بازده‌ها را مدل‌سازی کند و کاهش تدریجی نوسانات را نشان دهد.