

<https://chatgpt.com/share/69611f80-f550-800a-b6ae-d97eachbed5e2>

Me:

1 and 2 months lagged correlations

```
monthly_avg_precipitation_series_1 = monthly_avg_precipitation_series.shift(1)
```

```
monthly_avg_precipitation_series_2 = monthly_avg_precipitation_series.shift(2)
```

```
corr1_lag_p, p_value1_lag_p = pearsonr(monthly_avg_precipitation_series_1,  
all_cols_average)
```

```
print("Correlation:", corr1_lag_p, "p-value:", p_value1_lag_p)
```

```
corr1_lag_s, p_value1_lag_s = spearmanr(monthly_avg_precipitation_series_1,  
all_cols_average)
```

```
print("Correlation:", corr1_lag_s, "p-value:", p_value1_lag_s)
```

```
corr1_lag2_p, p_value1_lag2_p = pearsonr(monthly_avg_precipitation_series_2,  
all_cols_average)
```

```
print("Correlation:", corr1_lag2_p, "p-value:", p_value1_lag2_p)
```

```
corr1_lag2_s, p_value1_lag2_s = spearmanr(monthly_avg_precipitation_series_2,  
all_cols_average)
```

```
print("Correlation:", corr1_lag2_s, "p-value:", p_value1_lag2_s) böyle yaptım ama  
olmadı
```

Gemini:

Kodunun çalışmamasının temel sebebi büyük ihtimalle shift() işleminden sonra ortaya çıkan boş değerlerdir (NaN).

shift(1) yaptığında ilk satır boş kalır, shift(2) yaptığında ise ilk iki satır boş kalır. pearsonr ve spearmanr fonksiyonları (scipy kütüphanesinden geliyor) içinde NaN olan veri setleriyle işlem yapamaz ve hata verir ya da sonuç döndürmez.

Bunu çözmek için bu boş satırları temizlemen (drop) gerekiyor. İşte kodunun çalışan ve daha düzenli hali:

Çözüm: .dropna() Ekleme

Python

```
import pandas as pd
```

```
from scipy.stats import pearsonr, spearmanr
```

```
# 1. Verileri bir DataFrame'de topla ki hizalaması bozulmasın
```

```
data = pd.DataFrame({  
    'doluluk': all_cols_average,  
    'yagis_lag1': monthly_avg_precipitation_series.shift(1),  
    'yagis_lag2': monthly_avg_precipitation_series.shift(2)  
})
```

```
# 2. Shift işleminden dolayı oluşan boş satırları sil
```

```
# (Lag 2 kullandığın için ilk 2 satır silinecek)
```

```
data_clean = data.dropna()
```

```
# --- Lag 1 Analizi ---
```

```
corr1_p, p1_p = pearsonr(data_clean['yagis_lag1'], data_clean['doluluk'])
```

```
corr1_s, p1_s = spearmanr(data_clean['yagis_lag1'], data_clean['doluluk'])
```

```
print(f"Lag 1 - Pearson: {corr1_p:.3f} (p: {p1_p:.3e})")
```

```
print(f"Lag 1 - Spearman: {corr1_s:.3f} (p: {p1_s:.3e})")
```

```
# --- Lag 2 Analizi ---
```

```
corr2_p, p2_p = pearsonr(data_clean['yagis_lag2'], data_clean['doluluk'])
```

```
corr2_s, p2_s = spearmanr(data_clean['yagis_lag2'], data_clean['doluluk'])
```

```
print(f"\nLag 2 - Pearson: {corr2_p:.3f} (p: {p2_p:.3e})")
```

```
print(f"Lag 2 - Spearman: {corr2_s:.3f} (p: {p2_s:.3e})")
```

Neden Bu Hata Oldu?

- NaN Problemi: `shift(1)` yaptığında verin şöyle görünür: `[NaN, 10, 20, 30]`. İstatistiksel formüller "NaN" ile matematiksel işlem yapamaz.
- Boyut Uyuşmazlığı: Eğer `monthly_avg_precipitation_series_1` içindeki NaN değerini manuel silip `all_cols_average` serisindeki karşılığını silmezsen, bir seri diğerinden daha kısa olur ve fonksiyon hata verir. Yukarıdaki `dropna()` yöntemi her iki seriyi de aynı anda, karşılıklı olarak temizler.

Sonuçları Nasıl Yorumlamalısın?

Kodun çalıştıktan sonra şuna bak:

1. En yüksek Korelasyon: Hangi Lag (0, 1 veya 2) daha yüksek bir sayı veriyor? Genelde İstanbul barajları için Lag 1 veya Lag 2 daha yüksek çıkar.
2. P-Değeri: Eğer $p\text{-value} < 0.05$ ise sonucun anlamlıdır. $3.4e-05$ gibi bir şey görürsen bu çok küçük bir sayı demektir ($\$0.000034\$$), yani ilişkin çok güçlüdür.

Küçük bir tavsiye: Eğer hala hata alıyorsan, aldığın hatayı buraya kopyalayabilirsin. Büyük ihtimalle `shapes (X,) and (Y,) not aligned` (boyutlar uyuşmuyor) hatası alıyorsundur, yukarıdaki `data_clean` yöntemi bunu kesin çözecektir.

Bu çıktıları aldıktan sonra bunları metne dökmek ister misin? "Bulgular" kısmını yazmana yardımcı olabilirim.