Just analyzing my lazy running

In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as mdates
from datetime import datetime
```

In [2]:

```
df = pd.read_csv("running.csv")
df.head()
```

Out[2]:

	type	sourceName	sourceVersion	productType	device	startDate	endDate
0	HKWorkoutTypeIdentifier	Даня’s Apple Watch	8.5	Watch5,4	<<HKDevice: 0x28262be80>, name:Apple Watch,anu...	2022-03-12 03:37:35	2022-03-12 04:17:36
1	HKWorkoutTypeIdentifier	Даня’s Apple Watch	8.5	Watch5,4	<<HKDevice: 0x28262be80>, name:Apple Watch,anu...	2022-03-14 03:45:39	2022-03-14 04:25:45
2	HKWorkoutTypeIdentifier	Даня’s Apple Watch	8.5	Watch5,4	<<HKDevice: 0x28262be80>, name:Apple Watch,anu...	2022-03-15 03:43:31	2022-03-15 04:13:33
3	HKWorkoutTypeIdentifier	Даня’s Apple Watch	8.5	Watch5,4	<<HKDevice: 0x28262be80>, name:Apple Watch,anu...	2022-03-18 03:32:04	2022-03-18 03:59:07
4	HKWorkoutTypeIdentifier	Даня’s Apple Watch	8.5	Watch5,4	<<HKDevice: 0x28262be80>, name:Apple Watch,anu...	2022-03-19 03:35:12	2022-03-19 04:03:56

5 rows × 21 columns

In [3]:

```
cols = list(df.columns)
for i, col in enumerate(cols):
    print(f'{i+1}. {col}')
```

```
1. type
2. sourceName
3. sourceVersion
4. productType
5. device
6. startDate
7. endDate
8. activityType
9. duration
10. durationUnit
11. totalEnergyBurned
12. totalDistance
13. totalSwimmingStrokeCount
14. totalFlightsClimbed
15. HKElevationAscended
16. HKAverageMETs
17. HKWeatherHumidity
18. HKIndoorWorkout
19. HKTimeZone
```

```
20. HKWeatherTemperature
21. HKWasUserEntered
```

In [4]:

```
notNeeded = ['type', 'sourceName', 'sourceVersion', 'productType', 'device', 'activityTy

for col in notNeeded:
    cols.remove(col)
```

I wanted to see my '22 running year so I made a bar plot to realize how impermament I am...

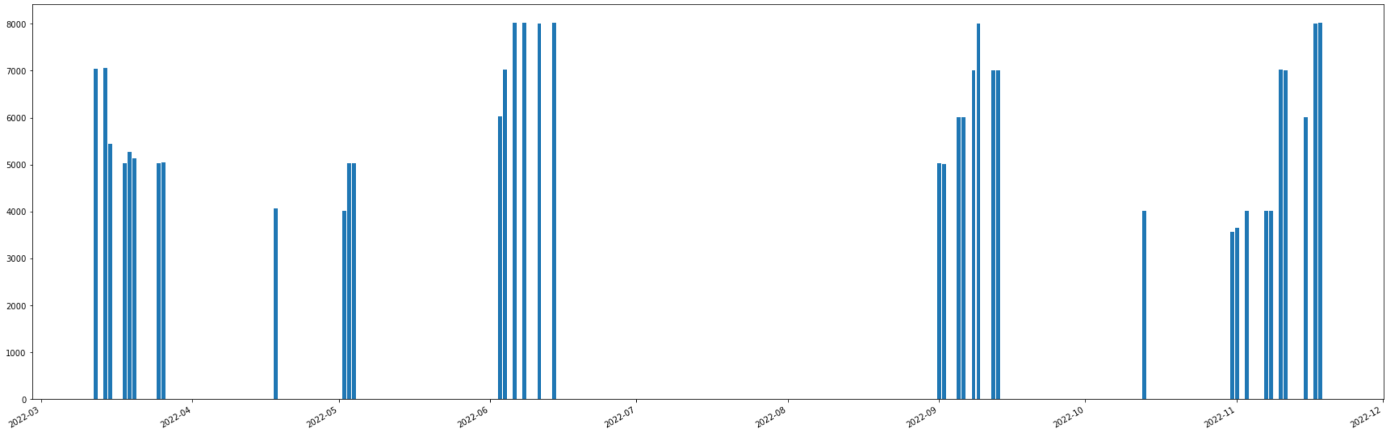
In [5]:

```
df['startDate'] = df['startDate'].map(lambda x: datetime.strptime(str(x), '%Y-%m-%d %H:%
x = df['startDate']
y = df['totalDistance']

plt.figure(figsize=(30,10))

# plot
plt.bar(x, y)
# beautify the x-labels
plt.gcf().autofmt_xdate()

plt.show()
```



And then I just calculated average time of my 8k runs

In [6]:

```
dur = df['duration']
dist = df['totalDistance']
total = 0.0
j = 0

for i, run in enumerate(dur):
    if dist[i] >= 8000.0:
        total += dur[i]/60
        j += 1

print(total/j)
```

```
42.182900883754094
```