```python
#Task 8: Use Windows with Structured Streaming to Track Most-Viewed URLs

from pyspark.sql import SparkSession
from pyspark.sql.functions import window, col

spark = SparkSession.builder.appName("Structured Streaming").getOrCreate()

# Simulated streaming data
data = [("url1", "2025-01-26 10:00:00"), ("url2", "2025-01-26 10:01:00")]
columns = ["URL", "Timestamp"]
static_df = spark.createDataFrame(data, columns)

# Process data with window
windowed_df = static_df.groupBy(window(col("Timestamp"), "10 minutes"), "URL").count()
windowed_df.show()

spark.stop()
```

```
+-------------------+----+-----+
|             window| URL|count|
+-------------------+----+-----+
|{2025-01-26 10:00...|url1|    1|
|{2025-01-26 10:00...|url2|    1|
+-------------------+----+-----+
```

+ Code   + Text

Start coding or generate with AI.