**[1] Probabilistic Forecasts of Bike-Sharing Systems for Journey Planning**[**https://dl.acm.org/doi/10.1145/2806416.2806569**](https://dl.acm.org/doi/10.1145/2806416.2806569)

1. It is not only importand, that there is a bike available at a station, but also that there is a free spot available when the user arrives at the destination station.
2. I don’t necceceraly want to know how many bikes are at a station at a given time, but the probability that a bike is available
3. Each station i = 1, … , N has a fixrd capacity ki, this is the maximum capacity, hence the maximum number of bikes that might be available at the station. The number of bikes currently available at station i at time t is noted as Xi(t) E {0, … , ki} and the number of availability for all station X at time t is X(t) E {X1(t) , …. , Xi(t)}. So each station has 2 states at the base level: if Xi(t) > 0, a bike for travel is available and Xi(t) < ki, a bike can be returned to that station. So, if a bike is taken the number of bikes decreases ( Xi(t) – 1 ) and if you bring a bike back to a station Xi(t) + 1.  
   What I will try to find out is the probability that a bike will be at a station at t + h, where h is the time added when the user needs the bike at a station, we also need to consider an available free spot at the destination station j so that all requirements Xi(t + h) > 0 and Xj((t+h) + traveltime) < kj are satisfied. This is, however, just the minimum requirements, the fluctuation of bikes also have to put into consideration so when the system gives the user a prediction, the number of bikes predicted to be available bikes or free spots should be higher than one. So the stochastic probability would be :

P(Ft | Xi(t+h) > 0 ^ Xj ( t + h + traveltime) < kj ))

1. Uses Markovian model, which means that I prediction is not based on the past data but the probability that a bike is available based on the current state of the stations. Those models have been succecfully tested in [7,10,12]. There is also the possibility of interconnecting the bike stations on a probabilistic basis in a way that the probability of a future time could be higher if bike stations nearby have currently more or less bikes, which indicates that in the close future some bikes will be brought back. This however would complicate the model immensely for very little gain as shown in

**[2]**

**[3]**

**[4]**

**[5]** **The-Bikeshare-Planning-Guide-ITDP:**

Dockless bikes are equipped with GPS so that the user does not need to find a specific station. This can however lead to the complication that in an area where there is not a high fluctuation of bikes you don’t find as many. With a fixed station you could potently have a more reliable source of available bikes then.

Several cities have made efforts to improve the ease and convenience of multi-modal trip

making by better integrating their bikeshare system with public transit. Operated by the transit

agency, Los Angeles Metro Bikeshare allows users to check out a bike using their Transit Access

Pass (TAP) card. Helsinki’s City Bikes system will be integrated into the mobility as a service

(MaaS) Whim app, which offers streamlined access to taxis, public transport, shared vehicles,  
and, soon, bikeshare through pay-as-you-go or monthly plans.1

**[6]** Modeling Bike Availability in a Bike-Sharing System Using Machine Learning

Spatio-temporal mobility model which combines the time and geographical components to predict bike availability.

The spatial dimension considers the geographical location and the relationship between thos locations. For example how the fluctuation on close stations usually interacts with each other.

### The time aspect examines time based changes on a daylie, weekly or monthly pattern and has to also involve special events. **Spatio-Temporal Mobility Model**

A spatio-temporal mobility model is a sophisticated analytical framework designed to predict and understand the movement patterns of entities (such as bikes in a bike-sharing system) across both spatial (geographical) and temporal (time-based) dimensions. This model integrates spatial and temporal data to capture the dynamic and complex behaviors of these entities over time and space.

**Key Components of a Spatio-Temporal Mobility Model**

1. **Spatial Data Analysis**:
   * **Geographical Locations**: The model considers the geographical positions of bike-sharing stations and other relevant locations within the urban environment.
   * **Spatial Relationships**: It analyzes the relationships and interactions between different locations, such as the proximity of stations to residential areas, business districts, and transit hubs.
2. **Temporal Data Analysis**:
   * **Time-Based Patterns**: The model examines time-based patterns, including daily, weekly, and seasonal variations in bike usage.
   * **Event Impact**: It considers the impact of special events, holidays, and weather conditions on bike availability and usage.
3. **Historical Data Utilization**:
   * **Usage Records**: By leveraging historical data, such as check-in and check-out records, the model identifies trends and patterns in bike usage.
   * **Predictive Insights**: This historical data is crucial for generating predictive insights about future bike availability and demand.
4. **Traffic Prediction Mechanism**:
   * **Flow Predictions**: The model forecasts the flow of bikes to and from each station, considering factors like the time of day, day of the week, and special events.
   * **Granular Predictions**: It provides predictions at a fine-grained level, often with sub-hour granularity, allowing for real-time adjustments.

**How Spatio-Temporal Mobility Models Work**

The spatio-temporal mobility model operates by integrating data from various sources to produce accurate predictions about bike availability and demand. Here’s how it typically works:

1. **Data Collection**: The model collects historical data on bike check-ins and check-outs from all stations. It also gathers external data, such as weather conditions, event schedules, and traffic patterns.
2. **Pattern Recognition**: Using advanced algorithms, the model identifies patterns in the data. For example, it might recognize that bike usage peaks in residential areas during the morning and in business districts during the evening.
3. **Prediction Generation**: The model uses these patterns to generate predictions about future bike availability. It can forecast which stations are likely to run out of bikes and which ones might have an excess at specific times.
4. **Rebalancing Strategies**: Based on these predictions, the model suggests rebalancing strategies. For instance, it might recommend moving bikes from low-demand stations to high-demand ones in anticipation of peak usage periods.

**Benefits of Spatio-Temporal Mobility Models**

* **Improved Accuracy**: By considering both spatial and temporal factors, the model provides more accurate predictions compared to traditional methods.
* **Dynamic Adaptation**: It allows for real-time adjustments and dynamic rebalancing, improving the overall efficiency of the bike-sharing system.
* **Enhanced User Experience**: With better predictions and rebalancing strategies, users are more likely to find bikes when and where they need them, enhancing satisfaction.
* **Operational Efficiency**: Operators can optimize their resources, reducing costs associated with bike redistribution and improving service levels.

**Example from Yang et al. (2016)**

Yang et al. (2016) proposed a spatio-temporal mobility model specifically for bike-sharing systems. Their model analyzed one year of data from a large bike-sharing system with over 2800 stations and 103 million check-in/out records. By incorporating both spatial and temporal dimensions, their model was able to predict bike availability with high accuracy. This prediction capability allowed operators to implement effective rebalancing strategies, ensuring that bikes were available at the right stations at the right times \cite{6}.

In conclusion, a spatio-temporal mobility model is a powerful tool for managing and optimizing bike-sharing systems. By leveraging comprehensive data analysis and predictive algorithms, it addresses the complex issue of bike distribution imbalance, enhancing both operational efficiency and user satisfaction.

**[7]**

**[8]**

**[9]   
Skellam regression:**

**Specifically models the difference between two count variables, z.b net total demand in bike sharing, which is the difference between bike rentals and returns.**

**Poisson variable is a variable which counts the number of events in a given time period, for example, how often does a bus arrive at a station in an hour or how many calls does a callcenter get in in hour. The numbers are always based on those counts and can are usually random.**

**[10]**

**[11]**

**[12]**

**[13]**

**[14]**

**[15]**

**[16]**

**[17]**

**[18]**

**[19]**

**[20]**

**From my Programming**

I started off with some sample data which seems like it was a snapshot from all stations at a given time : X(t). First I just printed out the data complete as a Json to get a genrell understanding of the structure of the data which was basically one big column of countries which included a list of nested dictionarys. My first thought was to get a good dataframe which I can use for my purpose. My first attempt was to set the indexes as countrynames and then all parameters as column names, since the citys and also the places inside those countries were also just nested dictionarys I changed the approach and extracted just the stations, since this is what I am aiming for anyway.