

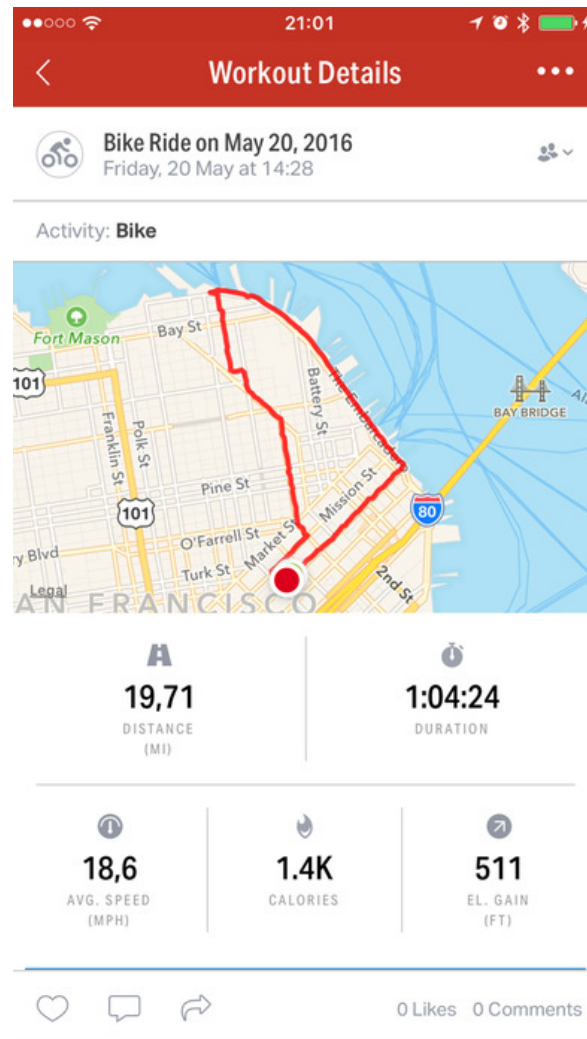
Project 1

16August 2015

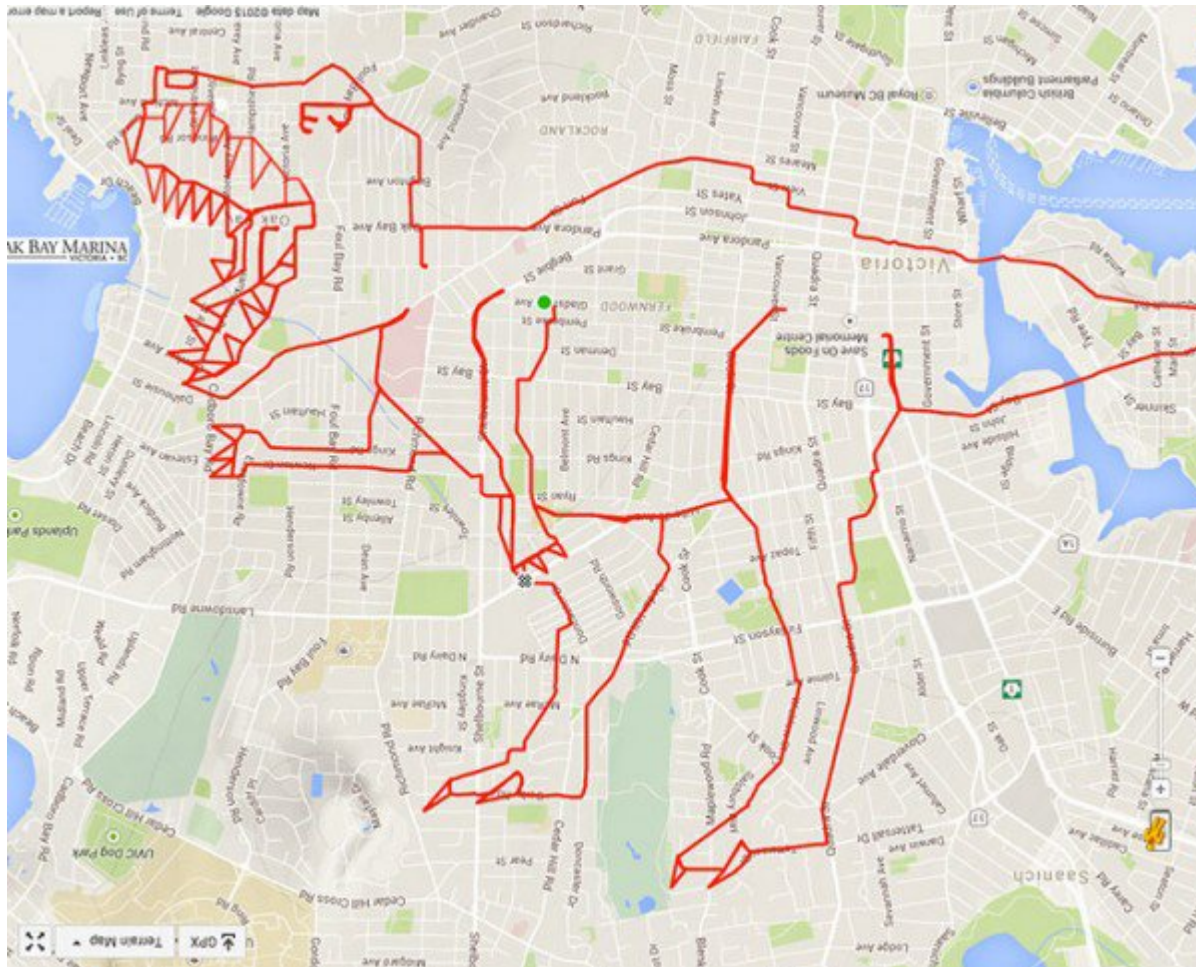
Data Science – Spatio-Temporal Data

- As Dr Salehi mentioned in her guest lecture – a major challenge is how to model complex spatio-temporal data sets
- Imagine you are a data science analyst for your company. You have GPS traces of the movement of your bicycle couriers. These GPS observations give the **trajectory** of the cyclist. What questions can you answer about their trajectories on different jobs?

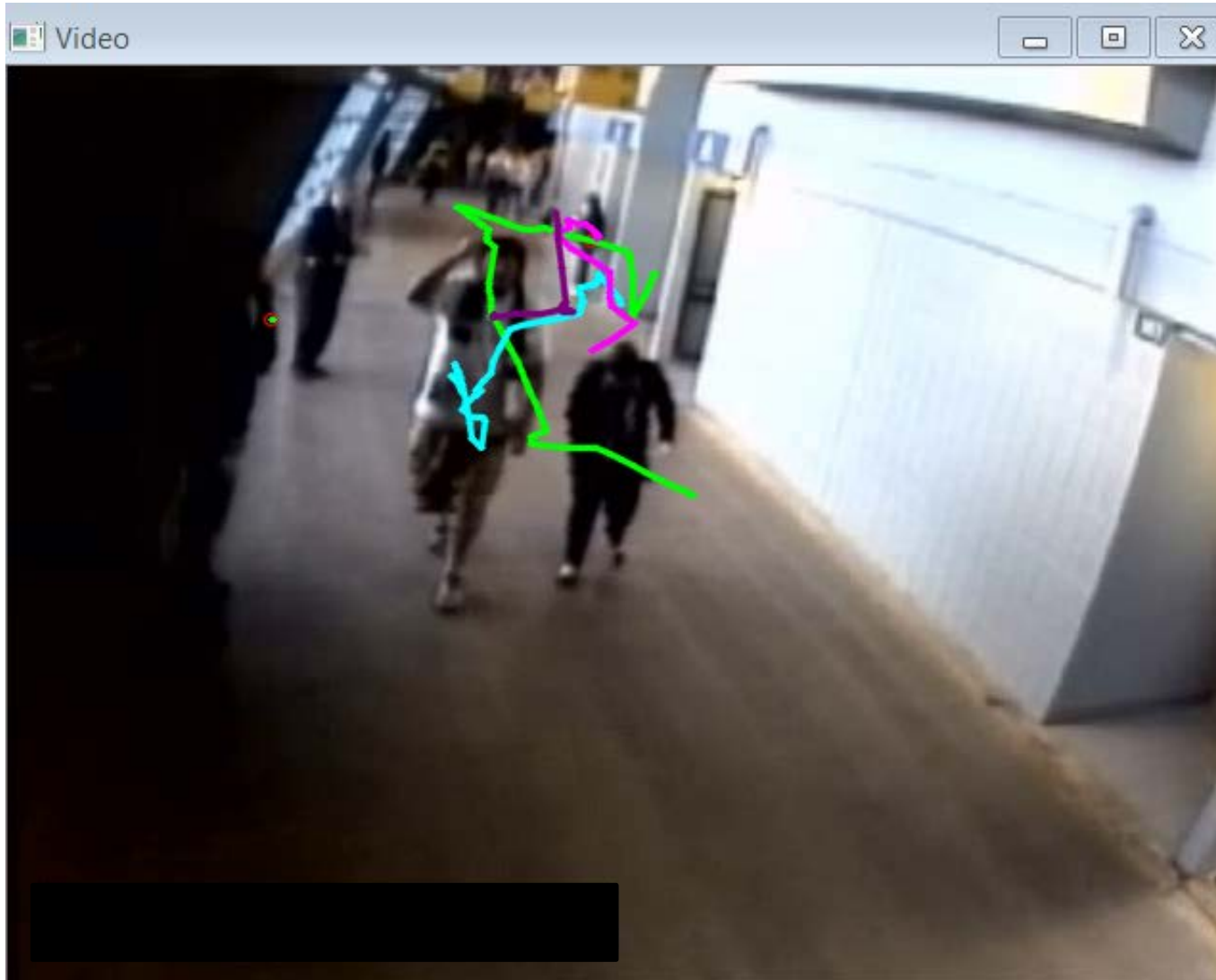
Example of GPS trace of cyclist I



Example of GPS trace of cyclist II



Example of trajectory extraction from video cameras I



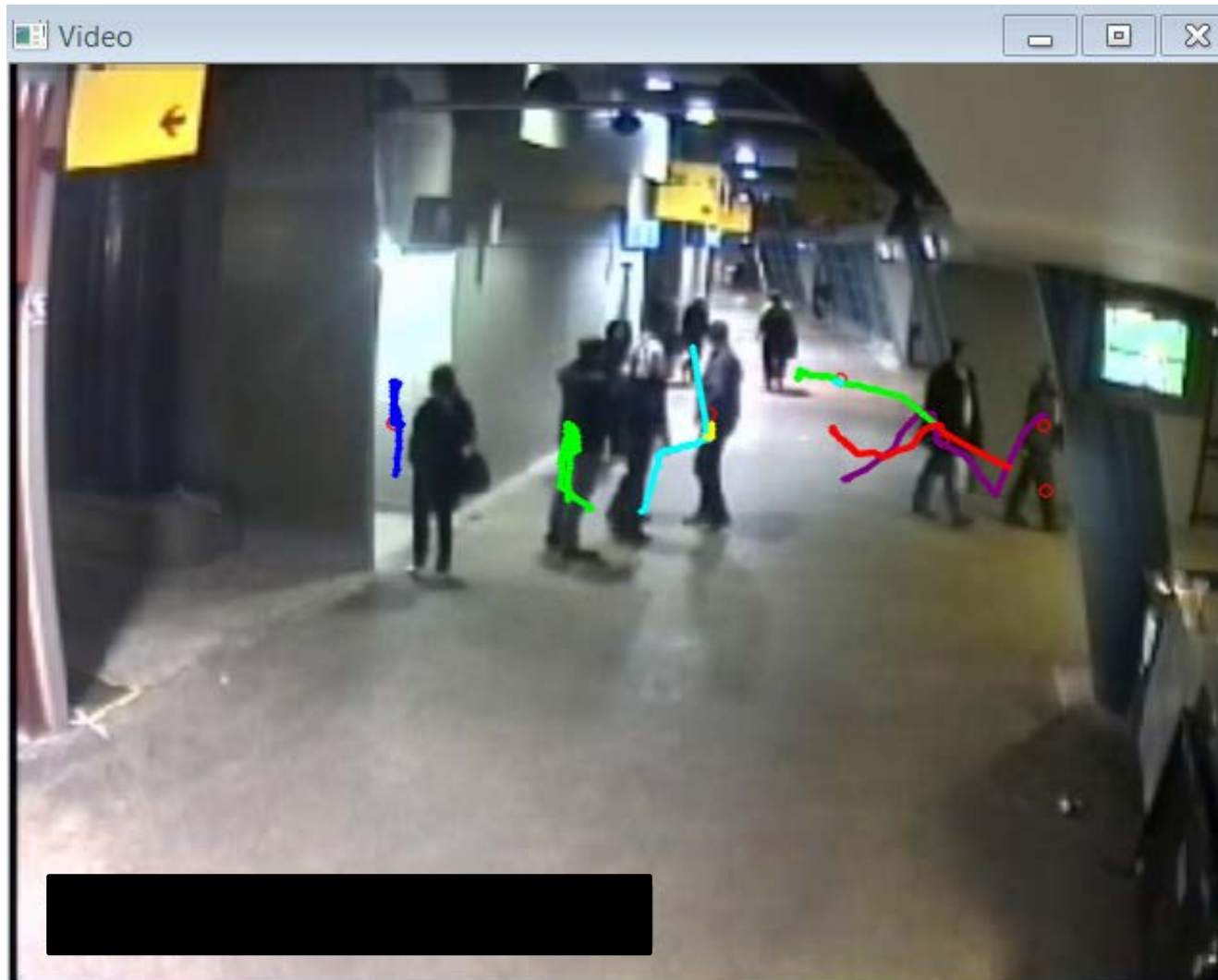
Thanks to
Meng Yang

Example of trajectory extraction from video cameras II



Thanks to
Meng Yang

Example of trajectory extraction from video cameras III

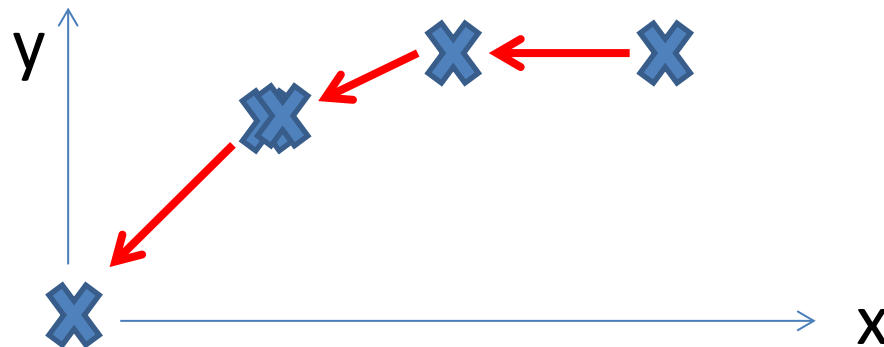


Thanks to
Meng Yang

Data Science – Trajectories

- We can represent a trajectory as a list of observations, where each observation is a tuple (x, y, t) containing the x, y coordinates of the object at time t

e.g. $[(3.0, 2.0, 1), (2.0, 2.0, 2), (1.0, 1.5, 4), (1.0, 1.5, 5), (0.0, 0.0, 6)]$



Data Science - Trajectories

- We will look at three different approaches to trajectory data mining
 - When is a trajectory observed to be closest to a given query location? (question 1)
 - How many times does a trajectory enter a given region? (question 2)
 - What is the longest period that two given trajectories are “close” in time and space? (question 3)

Question 1

You need to write a function that returns the time value of the closest observation in the trajectory to a given query point

You can assume that the observations in the trajectory are given in increasing time order

If there are multiple observations at the closest location, return the earliest one

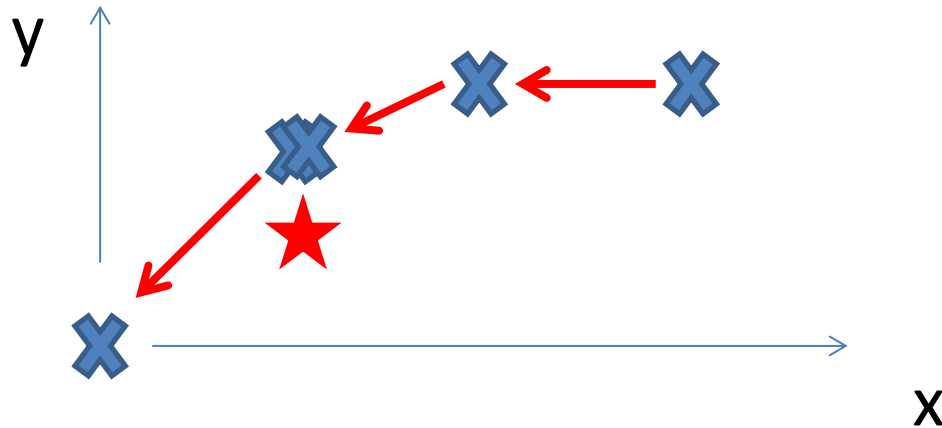
You should find the distance between two points (x_1, y_1) and (x_2, y_2) using the Euclidean distance:

$$d((x_1, y_1), (x_2, y_2)) = [(x_1 - x_2)^2 + (y_1 - y_2)^2]^{1/2}$$

Question 1

`closestTime([(3.0, 2.0, 1), (2.0, 2.0, 2), (1.0, 1.5, 4),
(1.0, 1.5, 5), (0.0, 0.0, 6)], (1, 1))`

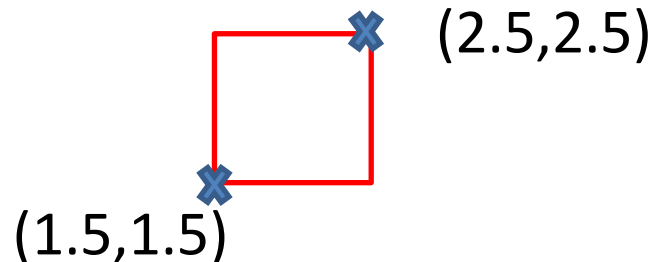
returns 4



Question 2

- You are given a trajectory and a query box
- How many times does the trajectory enter the query box?
- The query box is specified by its bottom-left and top-right coordinates

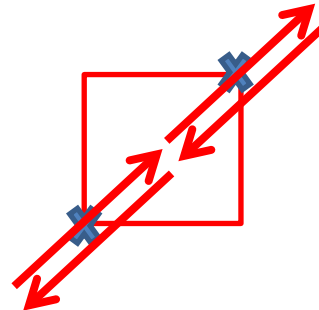
$((1.5, 1.5), (2.5, 2.5))$



Question 2

You need to write a function `visits(tr, b)` that counts the number of times the trajectory `tr` enters the box `b`

`visits([(1,1,0),(2,2,1),(3,3,2),(2,2,3),(1,1,4)], ((1.5,1.5),(2.5,2.5)))`
returns 2



Question 3

- Next, we want to check whether two objects are moving together
- What is the maximum number of consecutive observations in the two trajectories that are “close”?
- We can do this by comparing observations that occur at the same time in each trajectory, and checking if their distance is less than or equal to a threshold

Question 3

You need to write a function `together(tr1, tr2, th)` that takes two trajectories `tr1` and `tr2` and a distance threshold `th`, and returns the maximum number of consecutive observations such that the observations occur at the same time and are close.

```
together([(0,0,0),(1,1,1),(2,2,2),(1,1,3),(1,1,4)],  
          [(0,0,0),(6,1,1),(2,4,2),(1,1,3),(1,0.7,4)], 0.5)  
returns 2
```

```
together([(0,0,2),(1,1,3),(2,2,4),(1,1,5)],  
          [(0,0,0),(1,1,1),(2,2,2)], 0.5)  
returns 0
```

Academic Honesty

- All assessment items (worksheets, projects, test and exam) must be **your own, individual, original work**.
- For example, you must not copy the code of other students, and you must not make your code available to others to see. Do not give other students your login id and password, do not share USB memory drives, do not post your code on public forums, or any other activity that would make your code available to others. Likewise, do not ask other students to see their code. If other students ask to see your code, please say "no", as copying (collusion or plagiarism) is considered academic misconduct, and all students involved may face penalties (both the student who copied, and the student who made their code available).
- Any code that is submitted for assessment may be automatically compared against other students' code and other code sources using sophisticated similarity checking software, and cases of potential copying may lead to a formal academic misconduct hearing.
- For further information, please see the university's [Academic Honesty and Plagiarism](#) website, or ask your lecturer.

Conclusion

- The project questions will be submitted via Grok
- The specification of the project questions and the deadline will be announced very soon in Grok
- This project is worth 10% of the final subject
- We will be marking the correctness, quality, readability and commenting of your code