Building and veryfying the hypothesis

1. Predicting delivery time per sector

In order to predict delivery time per sector I would calculate the mean time of all deliveries, but for each sector individually. In the next step I would compare these values with the mean time of all deliveries. There is also a possibility to compare prediction errors for all deliveries end per sector.

2. An alternative method of predicting delivery time

In my opinion, to predict delivery time more accurately I would propose to group the data by the time of day (morning, afternoon, evening, night). Next, I would calculate the mean time for the different times of the day and different sectors and predict delivery times based on these groups.

3. Why could some deliveries take more time?

I think there is a wide range of reasons why some deliveries might take more time. Sometimes there could be an accident on the road and then the delivery time takes longer than usually. In the other case there could be traffic jams, especially in the morning or evening when most of people are commuting to school or work. Another reason could be a vehicle breakdown which would require some time to be repaired.

4. Collecting additional data

I think it would be worth collecting more data, for example information about the type of building, whether it's an apartment block or a house. Another valuable data would be the kind of transport a driver is using. If the driver uses a car, the delivery time will be shorter then when using a bicycle.

5. What is the risk of over- or under-estimating the delivery times?

I would say that the risk of overestimating the delivery time is a waste of time that could be used to deliver another order. On the other hand, underestimating the delivery time could cause that a customer can be unsatisfied because he must to wait longer for his order than he expected. This could lead to negative reviews about the delivery.