

## Assumptions and justifications

### D1.1:

- We decided to split the features column because there are columns such as colour, leglength etc. which were mostly empty and the values were in the features column. In the normalisation, as we have to deal with missing data, we thought it would be a good idea to fill these columns with the proper values and then check the missing data.
- We have named the entity order group as order.
- We have chosen a naming convention of lowercase and underscores with prefixes standing for the first letter of the full table name to improve readability and good practice for use in industry.
- We have added clustered indexes in the tables order and order\_item because are used in the Business Intelligence POC so are tables heavily queried.

### D1.2:

- We have made the assumption that only existing customers can place orders.
- We have also made the assumption that customers can't place orders for items that don't exist.
- We have also made the assumption that an order can't be placed if it already exists in the system. Same for the order item.

### D2:

- We found that we could create the cube without the need for creating a star/snowflake schema.