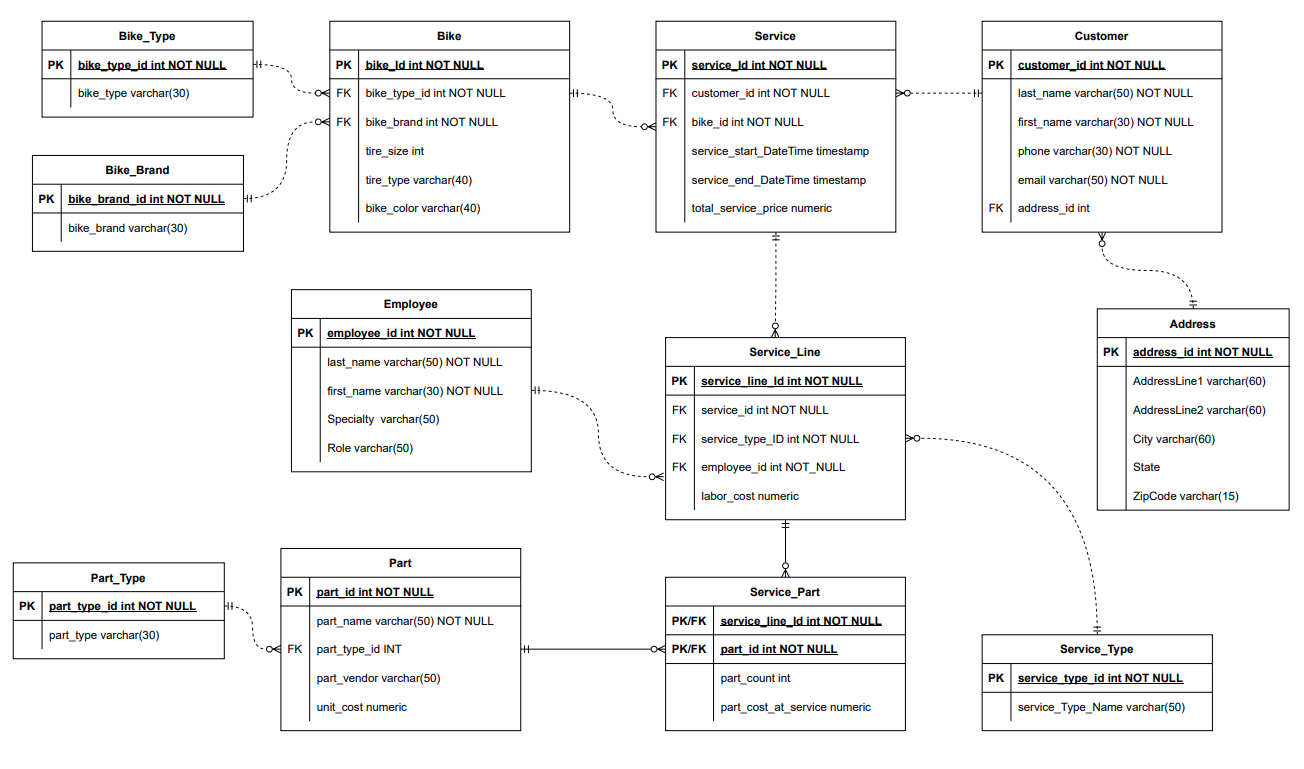
# Overview of the Assignment:

This assignment will go through steps to develop a data warehouse design.

# Part 1 – Review the business requirements

*BikeShop* is a national chain of bike repair shops which maintains a relational database to keep track of bike repairs. (example <https://www.landrys.com/> however consider them being more of a national business with hundreds of US locations)

*BikeShop* would like to introduce data warehousing and analytics to build their business. You have been hired as a data architect to create an initial Constellation data warehouse design. Below is *BikeShop’s* relational database. In addition, they would like to correlate the repair data with weather data to see if there are trends from weather which might impact repairs.

Here is a brief description of each of the tables in the OLTP system:

Customer and Address tables contain information about the customer. The address table is not normalized and may contain potential anomalies, especially if this data will need to be correlated with other external data.

For each bike, the type (i.e. mountain) and brand are tracked, along with tire size, type and bike color. Note that customers sometimes change the type of tire, and this data gets lost if a bike is brought back for service having a different tire type.

Service includes when the service started and ended, including the overall price. For each time the bike is serviced, there are multiple service lines (for example bike is brought in for tire and chain replacement which would be two service lines. An employee is tracked who performs the service. Some of the data around Employee is not normalized and contains anomalies. Each service line could have a part associated which includes prices for part as well as it’s cost at the time of service.

**External Weather data** is to come via an API such as <https://openweathermap.org> Here is a sample JSON API response to give you a sense of the data returned: <https://openweathermap.org/current#current_JSON>

{ "coord": { "lon": 10.99, "lat": 44.34 }, "weather": [ { "id": 501, "main": "Rain", "description": "moderate rain", "icon": "10d" } ], "base": "stations", "main": { "temp": 298.48, "feels\_like": 298.74, "temp\_min": 297.56, "temp\_max": 300.05, "pressure": 1015, "humidity": 64, "sea\_level": 1015, "grnd\_level": 933 }, "visibility": 10000, "wind": { "speed": 0.62, "deg": 349, "gust": 1.18 }, "rain": { "1h": 3.16 }, "clouds": { "all": 100 }, "dt": 1661870592, "sys": { "type": 2, "id": 2075663, "country": "IT", "sunrise": 1661834187, "sunset": 1661882248 }, "timezone": 7200, "id": 3163858, "name": "Zocca", "cod": 200 }

## Part 1 – Business Rules

1. Determine four business questions your data warehouse design will answer. Keep these questions in mind as you move on to the rest of the assignment. One of the questions needs to consider some sort of correlation with external weather data.

## Part 2 – Design a constellation schema warehouse

1. Create and insert an ERD showing the **constellation** ERD schema below. Please see requirements in the next two sections.
2. Determine four to five (non-date/time) dimension tables
   1. SCDs need to be considered- make sure to include at least one type 2 and one type 3
   2. Consider the date dimensions of different grains, consider a role-playing or bitemporal dimension, if so, explain why this is the case.
   3. For EACH table answer the following question: What kind of SCD is the dimension table and why did you choose this type of SCD?
3. Determine two to three fact tables
   1. How are the measures tied to your questions?
   2. Explain if your fact tables cumulative or snapshot
   3. Which attributes might be degenerative in your fact tables?
   4. Which attributes in the OLTP schema will transform to measures?
   5. What measures can be derived/calculated that should be included?

## Part 3 – Assumptions/Appendix (optional)

You may need to make some assumptions in order to work on this assignment. Clearly state any assumptions you needed to make here, along with your reasoning why the assumption was appropriate (one or two sentences should be sufficient in most cases). Do not remove any functionality, or trivialize, any of the assignment requirements. Feel free to augment.

Use the **Ask the Teaching Team Discussion Forum** if you have any questions regarding the how to approach this assignment.

Save your assignment as ***lastnameFirstname\_assign2\_0.docx*** and submit it in the *Assignments* section of the course.

For help uploading files please refer to the *Technical Support* page in the syllabus.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Criterion | A | B | C | D | F | Letter Grade |
| Technical mastery (50%) | Evidence of excellent mastery throughout | Evidence of good mastery throughout | Evidence of basic mastery throughout or good mastery intermittently | Minimal mastery evidenced | Virtually no mastery evidenced |  |
| Depth and thoroughness of coverage (25%) | Excellent depth and coverage of significant topics and issues | Good depth and coverage of significant topics and issues | Basic depth and coverage of significant topics and issues | Minimal depth and coverage of significant topics and issues | Virtually no depth and coverage of significant topics and issues |  |
| Clarity in presentation (25%) | Ideas and designs are exceptionally clear and organized throughout | Ideas and designs are clear and organized throughout | Ideas and designs are somewhat clear and organized throughout | Ideas and designs are mostly obscure and disorganized | Ideas and designs are entirely obscure and disorganized |  |
|  |  |  |  |  | Assignment Grade: |  |