Relational Schema

Business(
    Business_Id: varchar(25) [PK],
    name: varchar(30),
    address: varchar(65),
    city: varchar(20),
    state: varchar(15),
    postal_code: int [FK to Location.postal_code],
    latitude: float(10),
    longitude: float(10),
    stars: float(10),
    review_count: int,
    attributes: varchar(1000),
    categories: varchar(1000),
    monday_hours: varchar(9),
    tuesday_hours: varchar(9),
    wednesday_hours: varchar(9),
    thursday_hours: varchar(9),
    friday_hours: varchar(9),
    saturday_hours: varchar(9),
    sunday_hours: varchar(9)
)

Weather_By_Zip(
    postal_code: int [PK],
    month: varchar(9) [PK],
    avg_temp: float(20),
    high_temp: float(20),
    low_temp: float(20),
    total_rain_inches float(10)
)

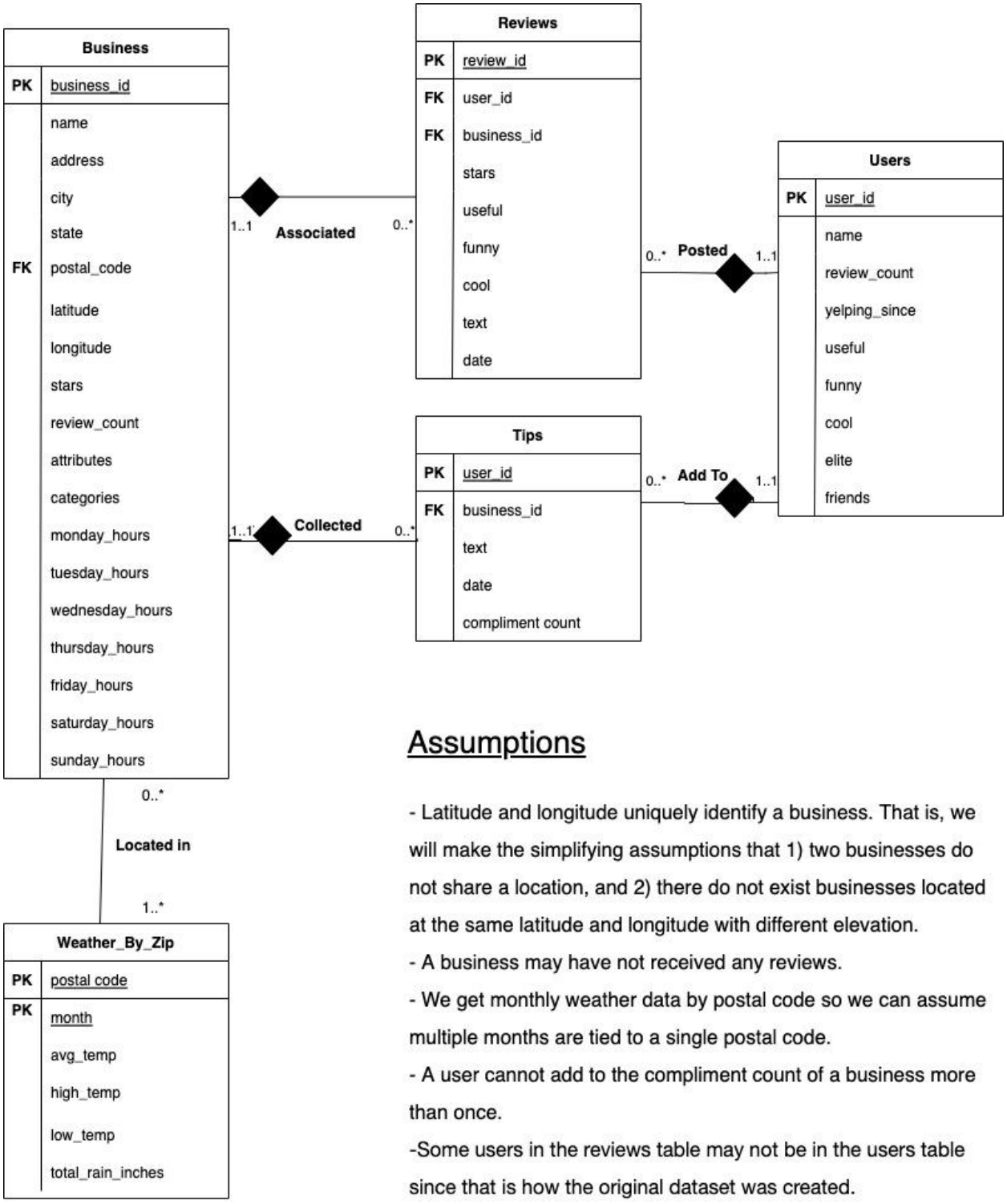
Reviews(
    review_id: varchar(25) [PK],
    user_id: varchar(25) [FK to Users.user_id],
    business_id: varchar(100) [FK to Business.Business_Id],
    stars: int,
    useful: int,
    funny: int,
    cool: int,
    text: varchar(6000),
    date: varchar(24)

)

Tips(
	tip_id: int [PK], user_id: varchar(25) [FK to Users.user_id],
	business_id: varchar(30),
	text: varchar(3000),
	date: varchar(24),
	compliment_count: int
)

Users(
	user_id: varchar(25) [PK],
	name: varchar(30),
	review_count: int,
	yelping_since: varchar(24),
	useful: int,
	funny: int,
	cool: int,
	elite: int,
	#don’t know how this is stored
	friends: int
)

ER/UML Diagram



Assumptions

- Latitude and longitude uniquely identify a business. That is, we will make the simplifying assumptions that 1) two businesses do not share a location, and 2) there do not exist businesses located at the same latitude and longitude with different elevation.
- A business may have not received any reviews.
- We get monthly weather data by postal code so we can assume multiple months are tied to a single postal code.
- A user cannot add to the compliment count of a business more than once.
-Some users in the reviews table may not be in the users table since that is how the original dataset was created.