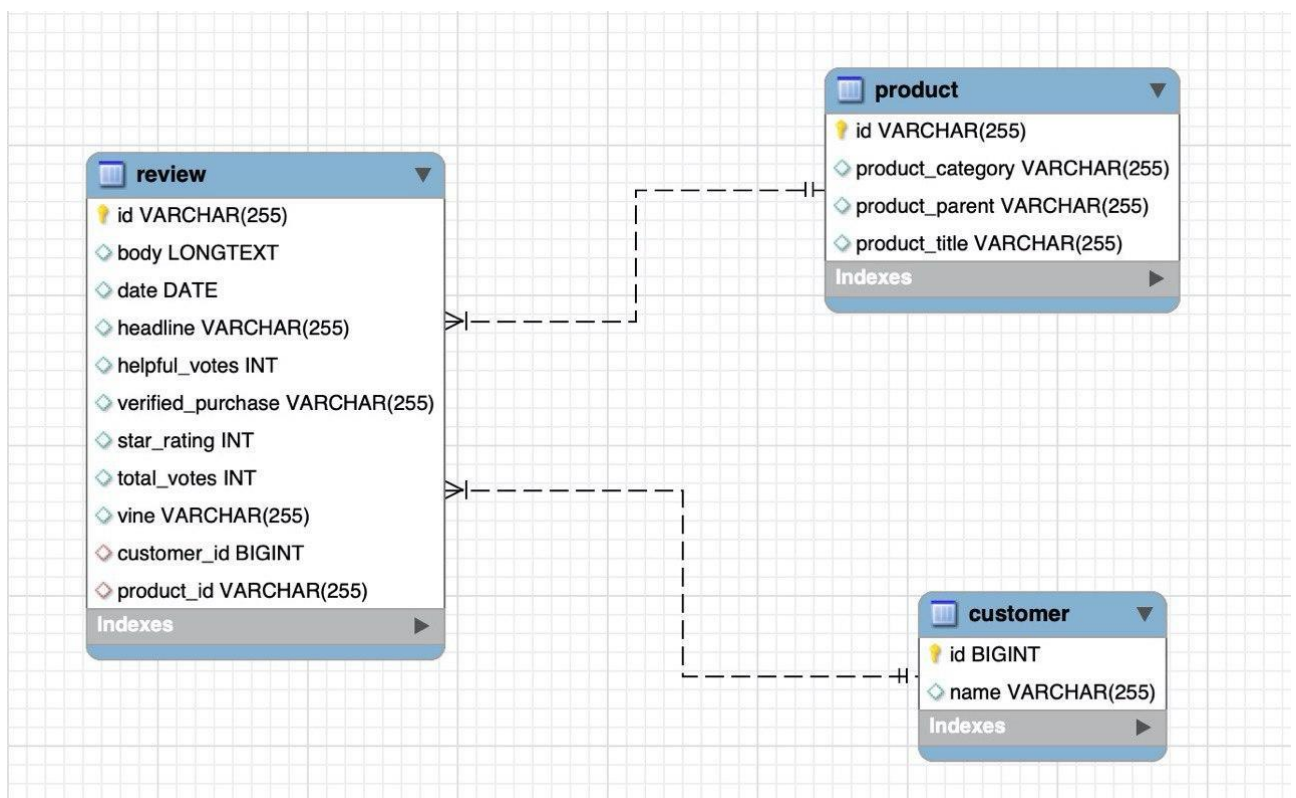# Questions

1. Return all reviews for specified `product_id`
2. Return all reviews for specified `product_id` with given `star_rating`
3. Return all reviews for specified `customer_id`
4. Return N most reviewed items (by # of reviews) for a given period of time
5. Return N most productive customers (by # of reviews written for verified purchases) for a given period
6. Return N best products by fraction of 5-star reviews (among products that have >100 of verified reviews)
7. Return N most productive "haters" (by # of 1- or 2-star reviews) for a given period
8. Return N most productive "backers" (by # of 4- or 5-star reviews) for a given period

RDBMS Diagram

product
- id VARCHAR(255)
- product_category VARCHAR(255)
- product_parent VARCHAR(255)
- product_title VARCHAR(255)
- Indexes

review
- id VARCHAR(255)
- body LONGTEXT
- date DATE
- headline VARCHAR(255)
- helpful_votes INT
- verified_purchase VARCHAR(255)
- star_rating INT
- total_votes INT
- vine VARCHAR(255)
- customer_id BIGINT
- product_id VARCHAR(255)
- Indexes

customer
- id BIGINT
- name VARCHAR(255)
- Indexes

Cassandra Diagram

Q1 Q2

product_reviews

rewiev_id	C
product_id	K
star_rating	C
helpful_votes	
total_votes	
vine	
verified_purchase	
review_headline	
review_body	
review_date	

Q3

customer_reviews

customer_id	K
review_id	
star_rating	C
helpful_votes	
total_votes	
vine	
verified_purchase	C
review_headline	
review_body	
review_date	C

Q6

best_products

product_parent	
product_category	
product_title	
product_id	K
verified_purchase	K

Q4 Q5 Q7 Q8

reviews_by_date_star

review_date	C
review_id	
product_parent	
product_category	
product_title	
product_id	K
verified_purchase	C
star_rating	C