**Mapping ER -> Relation**

USER

|  |  |
| --- | --- |
| userID | reviewerName |

PRODUCT

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| asin | category | duplicateRatio | incentivizedRatio | ratingAnamolyRate | reviewAnamolyRate |

REVIEW

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| reviewID | userID | asin | reviewText | minHash | overall | unixReviewTime | duplicate |

**Normalization**

USER

|  |  |
| --- | --- |
| userID | reviewerName |

FD1



PRODUCT

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| asin | category | duplicateRatio | incentivizedRatio | ratingAnamolyRate | reviewAnamolyRate |



FD1



REVIEW

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| reviewID | userID | asin | reviewText | minHash | overall | unixReviewTime | duplicate | incentivized |



FD1



**1NF**

1. Is each relation flat?
   1. Yes
2. Are there only single, atomic, indivisible attributes?
   1. Yes

**Relations are in 1NF**

**2NF**

1. Are there functional dependencies? (B is functionally dependent on A if **each value of A** in relation R is associated with only **one value of B** at **any time**.

**Rules:**

* For each set of attributes on the left side of the FD, there is only one value to the right
* This rule must hold at all times
* FD must be nontrivial

**Assumptions:**

* There can be multiple reviews written at the same time
* Two reviews can have the same text and minHash (those would be duplicate reviews)
* A reviewID depends on the product a user writes a review for.
* {reviewText} -> {minHash} is nontrivial because minHash is not a subset of reviewText, but rather a hash found through calculation.
* Two reviews may be identical, so they may have identical minHashes (a review text may not be unique). Therefore, the functional dependency {reviewText, minHash} -> {duplicate, incentivized} does not hold.

**Reviews**

Functional dependencies:

* {reviewID} -> {reviewText, minHash, overall, unixReviewTime, duplicate, incentivized}
  + Key -> non-key
* Can I remove anything from the left side and still have a FD hold?
  + No because review is a weak entity type, and therefore requires userID and asin
* Primary Key = {reviewID}

**Users**

Functional dependencies:

* {userID} -> {reviewerName}
  + Key -> non-key
* Can I remove anything from the left side and still have a FD hold?
  + No because review is a weak entity type, and therefore requires userID and asin
* Primary Key = {reviewID, userID, asin}

**Products**

Functional dependencies:

* {asin} -> {category, duplicateRatio, incentivizedRatio, ratingAnamolyRate, reviewAnamolyRate}
  + Key -> non-key
* Can I remove anything from the left side and still have a FD hold?
  + No because review is a weak entity type, and therefore requires userID and asin
* Primary Key = {reviewID, userID, asin}

1. Decompose

**Reviews**

N/A

**Users**

N/A

**Products**

N/A

1. Are all FD’s Full functional dependencies?

**Yes,** **all relations are in 2NF**

**3NF**

There are no non-key -> non-key FD’s and no transitive FD’s.

**All relations are in 3NF**