EduRank STEP 3

Group 27 Members:

Hayrettin Çağan Şendağ 27013

Oğul Sümer 27941

Efe Temur 29061

Yiğit Kaan Tonkaz 29154

Github Repository:

https://github.com/hayrettins/EduRank

The writing in red between the // symbols are explanations of SQL statements. The SQL statements are written in black

```
CREATE OR REPLACE VIEW LITERACY_BY_YEAR
AS
SELECT report_id,country_code , rate, year
FROM literacy_reports;

SELECT * FROM LITERACY_BY_YEAR;
```

//

LITERACY_BY_YEAR allows us to create a view that presents the increase in literacy at each country by year. In this view we are free to observe how many years have passed since the last report and how much the literacy rate has increased in those years. We created this view because it also allows us to calculate an average increase value in literacy by year for each country. Suddenly we are able to decide which country put in more work to increase literacy using actual values.

//

```
CREATE OR REPLACE VIEW LITERACY_BY_AGE
AS
SELECT report_id, country_code, rate, age
FROM literacy_reports;

SELECT * FROM LITERACY_BY_AGE;
```

//

LITERACY_BY_AGE allows us to observe how in each report instance a countries literacy rate increased in adult and young groups. We created this view in order to observe and calculate how young and adult groups growth rate of literacy rate changes in each country. A country with more increase in young groups would be a factor supporting the idea that that country has a better quality of teaching.

//

```
CREATE OR REPLACE VIEW LITERACY_BY_GENDER
AS
SELECT report_id, country_code, rate, gender
FROM literacy_reports;
```

```
SELECT * FROM LITERACY_BY_GENDER;
//
```

LITERACY_BY_GENDER allows us to observe the growth rate of each sex in each country while also helping us gather actual data that will contribute to the teaching quality of the country later on. A higher percentage of female literacy would mean a lot more since it is known that there are several countries that still refuses to contribute to the education of women.

```
//


CREATE OR REPLACE VIEW LITERACY_ANALYSIS
AS
SELECT COUNTRY_CODE, COUNT(REPORT_ID) REPORT_COUNT,
MAX(RATE) MAX_RATE, MIN(RATE) MIN_RATE, AVG(RATE) AVG_RATE
FROM LITERACY_REPORTS
GROUP BY COUNTRY_CODE;

SELECT * FROM LITERACY_ANALYSIS;

//
```

In LITERACY_ANALYSIS we aimed to calculate the min, max, average, values of literacy rate while keeping count of the number of reports on each country. A higher report count would mean that we have more accurate data on that country. The highest max value would mean a higher level of education and we can isolate those countries to observe other factors that contributed to this outcome later on. We have observed them by sex and age. We are able to see that higher percentage of women literacy correlates to higher percentage of literacy in general. We could also say that segregation by sex in education means less improvement in teaching and literacy for such countries. The min value also helps support this claim. The average values allow us to make a statement about how they changed over time. We can see which country realized the importance of education and tried to make a change about it.

```
//


CREATE OR REPLACE VIEW LIT_EXP_V
AS
SELECT L.report_id,L.COUNTRY_CODE, L.RATE, L.YEAR, L.GENDER, L.AGE, E.exp_percentage
FROM literacy_reports AS L
JOIN  expenditures AS E ON(L.country_code = E.country_code AND L.YEAR = E.YEAR);

SELECT * FROM LIT_EXP_V;
```

//

In LIT_EXP_V we can see the correlation between the money spent on education and change in literacy rate. The outcome is the same with our thesis which is , countries who spend more of their budget on education have reached higher level of literacy in young and female groups(and in general) compared to the other countries.

//

```
CREATE OR REPLACE VIEW LIT_ENR_V
AS
SELECT L.report_id,L.COUNTRY_CODE, L.RATE, L.YEAR, L.GENDER, L.AGE, E.ratio
FROM literacy_reports AS L
JOIN  enrollment AS E ON(L.country_code = E.country_code AND L.YEAR = E.YEAR);

SELECT * FROM LIT_ENR_V;
```

//

In LIT_ENR_V we can observe the correlation between enrollment rate and literacy rate. This view supports the idea that literacy rate and enrollment rate are directly proportional. The gender data helps us see which countries have low enrollment rate for girls by showing the increase in literacy rate and the enrollment rate. If the enrollment rate is high but literacy rate does not experience a steep rise we could predict that this country does not allow the enrollment of girls(Afghanistan). The same math could be done with the age data to show the countries that did not make change for the better.

//

```
CREATE OR REPLACE VIEW LIT_QUAL_V
AS
SELECT L.report_id, L.COUNTRY_CODE, L.RATE, L.YEAR, L.GENDER, L.AGE, E.percentage
FROM literacy_reports AS L
JOIN  teacher_quality AS E ON(L.country_code = E.country_code AND L.YEAR = E.YEAR);

SELECT * FROM LIT_QUAL_V;
```

//

LIT_QUAL_V view helps us observe the correlation between teacher percentage and literacy rate while showing us other data(age, gender, country) that will help us make other hypothesis. We can see the percentage of teacher and literacy rate are directly proportional. When we zoom in on the gender data we can see that this rule does not apply to females and adult on some countries. Which will be factors to decrease their education quality later on.

```
//
SELECT L.age, L.country_code, AVG(L.rate) AS avg_literacy_rate, MIN(L.rate) AS
min_literacy_rate,
COUNT(L.report_id) AS literacy_report_count, AVG(E.exp_percentage) AS avg_expenditure
FROM literacy_reports L
JOIN expenditure_report E ON L.country_code = E.country_code AND L.year = E.year
GROUP BY L.age, L.country_code
HAVING AVG(L.rate) > 0.7 AND COUNT(L.report_id) >= 10;

//
```

In this query we are able to observe the average and min literacy rates for each age group and country with their average expenditure on education. We check the literacy report count to ensure there are sufficient reports available for analysis. We find that countries with the lowest rate for adult literacy has the highest expenditure rate. We can conclude that such countries are trying to make an effort in order to increase their adult literacy rate. /*average and minimum literacy rates for each age group and country with their average expenditures on education*/

```
//


SELECT L.gender,L.country_code, AVG(E.exp_percentage) AS avg_expenditure,
MAX(L.rate) AS max_literacy_rate, MIN(L.rate) AS min_literacy_rate, AVG(L.rate) AS
avg_literacy_rate, COUNT(L.report_id) AS literacy_report_count
FROM literacy_reports L
JOIN expenditure_report E ON L.country_code = E.country_code AND L.year = E.year
GROUP BY L.gender, L.country_code
HAVING COUNT(L.report_id) >= 5 AND AVG(L.rate) >= 90;



//
```

The objective of the query is to identify countries and genders where the literacy rate is relatively high, and check the literacy report count to ensure there are sufficient reports available for analysis. We can observe the male and female average literacy rates for the same country which allows us to come to a conclusion about that countries education quality. We can also come to the conclusion that, expenditure rate for education is insignificant while considering the difference between average literacy for male and female , for most countries.

The query also filters out countries and genders with a low literacy report count or a low average literacy rate to ensure that the data is statistically significant.

//

```
SELECT E.country_code, AVG(L.rate) AS avg_literacy_rate,
COUNT(L.report_id) AS literacy_report_count, AVG(E.exp_percentage) AS avg_expenditure
FROM expenditure_report E
JOIN literacy_reports L ON E.country_code = L.country_code AND E.year = L.year
JOIN enrollment_report ER ON E.country_code = ER.country_code AND E.year = ER.year
WHERE ER.ratio > 0.9
GROUP BY E.country_code
HAVING COUNT(L.report_id) >= 5;
```

//

The result of this query gives us the average expenditure on education by country, for years where the enrollment ratio is above 90%. We check the literacy report count to ensure there are sufficient reports available for analysis. 9 results are returned to us and we can see that average literacy rate and expenditure are directly proportional where enrollment rate is over 90 percent.

//

PROS AND CONS:

Creating general constraints is relatively easy and efficient as they are integrated with the database engine and optimized for performance. But they can't handle exceptions. Triggers provide flexibility in handling exceptions to the validation rules. But, creating and maintaining triggers are difficult because they require writing custom code, and they may negatively impact database performance if they are written badly. Also, triggers can be disabled, which may lead to compromised data integrity.


```
SELECT MIN(rate), MAX(rate) FROM literacy_reports;
```
// selects the minimum and maximum values from the "rate" column of the "literacy_reports" table//
```
ALTER TABLE literacy_reports ADD CONSTRAINT rate_range CHECK (rate >= 3.39 AND rate <= 100) ;
```
// adds a constraint to the "literacy_reports" table that ensures that the "rate" column can only have values between 3.39 and 100//
```
INSERT INTO literacy_reports (report_id, country_code, rate, year, gender, age)
VALUES (3351, 'TUR', 120, 2022, 'male', 'young');
```
// insert a new row into the "literacy_reports" table with a "rate" value of 120, which violates the newly added constraint//
```
DROP TRIGGER IF EXISTS fix_rate_before_insert;
DROP TRIGGER IF EXISTS fix_rate_before_update;
```
// drop any existing triggers named "fix_rate_before_insert" and "fix_rate_before_update", respectively//


```
DELIMITER //
CREATE TRIGGER fix_rate_before_insert
BEFORE INSERT ON literacy_reports
FOR EACH ROW
BEGIN
   IF NEW.rate > 100 THEN
      SET NEW.rate = 100;
   ELSEIF NEW.rate < 3.39 THEN
      SET NEW.rate = 3.39;
   END IF;
END//
DELIMITER ;
```
// creates a new trigger named "fix_rate_before_insert" that runs before each insert on the "literacy_reports" table. The trigger checks if the new "rate" value is outside the allowed range (3.39 to 100) and sets it to the appropriate value if necessars//

```sql
DELIMITER //
CREATE TRIGGER fix_rate_before_update
BEFORE UPDATE ON literacy_reports
FOR EACH ROW
BEGIN
   IF NEW.rate > 100 THEN
      SET NEW.rate = 100;
   ELSEIF NEW.rate < 3.39 THEN
      SET NEW.rate = 3.39;
   END IF;
END//
DELIMITER ;
```
// creates a new trigger named "fix_rate_before_update" that runs before each update on the "literacy_reports" table. The trigger checks if the updated "rate" value is outside the allowed range (3.39 to 100) and sets it to the appropriate value if necessary//

```sql
INSERT INTO literacy_reports (report_id, country_code, rate, year, gender, age)
VALUES (3351, 'TUR', 120, 2022, 'male', 'young');
```
// attempts to insert a new row into the "literacy_reports" table with a "rate" value of 120, but the trigger from the sixth SQL statement changes the value to 100//
```sql
SELECT * FROM literacy_reports where report_id = 3351;
```
// selects all columns from the "literacy_reports" table where the "report_id" is 3351. At this point, there should be only one row with that ID//

```sql
DELETE FROM literacy_reports WHERE report_id = 3351;
```
// deletes the row from the "literacy_reports" table where the "report_id" is 3351/

==Stored Procedure:==
```sql
DROP PROCEDURE IF EXISTS get_lit_stats;

DELIMITER //
CREATE PROCEDURE get_lit_stats(IN iso_code VARCHAR(3))
BEGIN
        DECLARE code_exists INT DEFAULT 0;
   SELECT COUNT(*) INTO code_exists FROM literacy_reports WHERE country_code = iso_code;
   IF code_exists > 0 THEN
      SELECT * FROM literacy_reports WHERE country_code = iso_code;
   ELSE
      SIGNAL SQLSTATE '45000' SET MESSAGE_TEXT = 'Country code does not exist in the
literacy_reports table.';
   END IF;
END//
```

```
DELIMITER ;

CALL get_lit_stats('TUR');
CALL get_lit_stats('AFG');

CALL get_lit_stats('TDR');

//
```

```
//
```