**Google BigQuery commands and Results**

-- Using Google big query from Google Cloud Platform

-- 2.1 Extract all data to ensure the data load

SELECT \*

FROM

    `ace-technical-test-q2.ffb\_yield.ffb`

LIMIT 1000

-- 2.2 check any missing values

SELECT

    COUNT(\*) AS Missing\_Values

FROM

    `ace-technical-test-q2.ffb\_yield.ffb`

WHERE

    SoilMoisture IS NULL

    OR Average\_Temp IS NULL

    OR Min\_Temp IS NULL

    OR Max\_Temp IS NULL

    OR Precipitation IS NULL

    OR Working\_days IS NULL

    OR HA\_Harvested IS NULL

OR FFB\_Yield IS NULL;

|  |
| --- |
| Result:  [{    "Missing\_Values": "0"  }] |

-- 2.3 group data by year and month

--perform calculation for each external factors

SELECT

    EXTRACT(YEAR FROM PARSE\_DATE('%d.%m.%Y', Date)) AS Year, /\* change date format \*/

    EXTRACT(MONTH FROM PARSE\_DATE('%d.%m.%Y', Date)) AS Month,

    AVG(SoilMoisture) AS Avg\_SoilMoisture,

    AVG(Average\_Temp) AS Avg\_Average\_Temp,

    AVG(Min\_Temp) AS Avg\_Min\_Temp,

    AVG(Max\_Temp) AS Avg\_Max\_Temp,

    AVG(Precipitation) AS Avg\_Precipitation,

    SUM(Working\_days) AS Total\_Working\_days,

    SUM(HA\_Harvested) AS Total\_HA\_Harvested,

    AVG(FFB\_Yield) AS Avg\_FFB\_Yield

FROM

    `ace-technical-test-q2.ffb\_yield.ffb`

GROUP BY

    Year,

    Month

ORDER BY

    Year,

    Month;

-- 2.4 obtain the summary statistics based on the aggregated data

SELECT

    AVG(SoilMoisture) AS Avg\_SoilMoisture,

    MIN(SoilMoisture) AS Min\_SoilMoisture,

    MAX(SoilMoisture) AS Max\_SoilMoisture,

    AVG(Average\_Temp) AS Avg\_Average\_Temp,

    MIN(Average\_Temp) AS Min\_Average\_Temp,

    MAX(Average\_Temp) AS Max\_Average\_Temp,

    AVG(Precipitation) AS Avg\_Precipitation,

    MIN(Precipitation) AS Min\_Precipitation,

    MAX(Precipitation) AS Max\_Precipitation,

    AVG(FFB\_Yield) AS Avg\_FFB\_Yield

FROM

    `ace-technical-test-q2.ffb\_yield.ffb`;

|  |
| --- |
| Result:  [{    "Avg\_SoilMoisture": "527.646923076923",    "Min\_SoilMoisture": "380.7",    "Max\_SoilMoisture": "647.3",    "Avg\_Average\_Temp": "26.849918480923062",    "Min\_Average\_Temp": "25.15806452",    "Max\_Average\_Temp": "28.58",    "Avg\_Precipitation": "188.98076923076923",    "Min\_Precipitation": "2.0",    "Max\_Precipitation": "496.1",    "Avg\_FFB\_Yield": "1.6022307692307693"  }] |

--2.5 Correlation

SELECT

    CORR(SoilMoisture, FFB\_Yield) AS Corr\_SoilMoisture\_FFB\_Yield,

    CORR(Average\_Temp, FFB\_Yield) AS Corr\_Average\_Temp\_FFB\_Yield,

    CORR(Precipitation, FFB\_Yield) AS Corr\_Precipitation\_FFB\_Yield,

    CORR(Working\_days, FFB\_Yield) AS Corr\_Workingdays\_FFB\_Yield,

    CORR(HA\_Harvested, FFB\_Yield) AS Corr\_Harvested\_FFB\_Yield,

FROM

    `ace-technical-test-q2.ffb\_yield.ffb`;

|  |
| --- |
| Result:  [{    "Corr\_SoilMoisture\_FFB\_Yield": "-0.0031829013542888328",    "Corr\_Average\_Temp\_FFB\_Yield": "-0.0054943529235982856",    "Corr\_Precipitation\_FFB\_Yield": "0.28960372412286867",    "Corr\_Workingdays\_FFB\_Yield": "0.116364072327534",    "Corr\_Harvested\_FFB\_Yield": "-0.35022183838334348"  }] |

--2.6 trend analysis in year for FFB yield, precipitation and area harvested

SELECT

    EXTRACT(YEAR FROM PARSE\_DATE('%d.%m.%Y', Date)) AS Year,

    AVG(FFB\_Yield) AS Avg\_FFB\_Yield,

    AVG(Precipitation) AS Avg\_Precipitation,

    AVG(HA\_Harvested) AS Avg\_Harvested\_Area

FROM

    `ace-technical-test-q2.ffb\_yield.ffb`

GROUP BY

    Year

ORDER BY

    Year;

|  |
| --- |
| Result:  [{    "Year": "2008",    "Avg\_FFB\_Yield": "1.6125",    "Avg\_Precipitation": "206.44166666666666",    "Avg\_Harvested\_Area": "787123.80468333326"  }, {    "Year": "2009",    "Avg\_FFB\_Yield": "1.5733333333333333",    "Avg\_Precipitation": "182.21666666666664",    "Avg\_Harvested\_Area": "794979.47335833323"  }, {    "Year": "2010",    "Avg\_FFB\_Yield": "1.5",    "Avg\_Precipitation": "184.55833333333331",    "Avg\_Harvested\_Area": "807303.16872500011"  }, {    "Year": "2011",    "Avg\_FFB\_Yield": "1.6458333333333333",    "Avg\_Precipitation": "223.54166666666666",    "Avg\_Harvested\_Area": "786768.65165833349"  }, {    "Year": "2012",    "Avg\_FFB\_Yield": "1.585",    "Avg\_Precipitation": "189.99166666666665",    "Avg\_Harvested\_Area": "802893.68623333343"  }, {    "Year": "2013",    "Avg\_FFB\_Yield": "1.6241666666666665",    "Avg\_Precipitation": "202.95833333333331",    "Avg\_Harvested\_Area": "792260.86085"  }, {    "Year": "2014",    "Avg\_FFB\_Yield": "1.625",    "Avg\_Precipitation": "174.05",    "Avg\_Harvested\_Area": "780182.88943333342"  }, {    "Year": "2015",    "Avg\_FFB\_Yield": "1.6666666666666665",    "Avg\_Precipitation": "155.17499999999998",    "Avg\_Harvested\_Area": "775253.31668333337"  }, {    "Year": "2016",    "Avg\_FFB\_Yield": "1.4841666666666669",    "Avg\_Precipitation": "162.23333333333332",    "Avg\_Harvested\_Area": "782140.941925"  }, {    "Year": "2017",    "Avg\_FFB\_Yield": "1.7241666666666668",    "Avg\_Precipitation": "217.69166666666663",    "Avg\_Harvested\_Area": "800756.45623333321"  }, {    "Year": "2018",    "Avg\_FFB\_Yield": "1.58",    "Avg\_Precipitation": "178.12",    "Avg\_Harvested\_Area": "822662.49061"  }] |