

Analysis of Lifestyle and well-being Data Using Bigquery and Power BI

By ONwadinobi
February 7, 2024.

Background

The health behavior of individuals has implications on how healthy or not individuals are, and to what extent they are able to optimize their well-being for overall optimum health and productivity.

Being a professional health care practitioner, and particularly in the area of emergency care, with keen interest in appraising the health behavior of both health care practitioners and those who utilize health care services to appreciate insights aimed at contributing to finding ways to educate and encourage people towards engaging in positive health behavior.

Objective

With the persisting pressure of workload facing health care workers in health facilities and the waiting time for patients to be attended to, and also understanding the impact stress has on quality of sleep, physical activity, body mass index(BMI), daily productivity and general well-being, these encouraged me to find data on lifestyle and behavior such as this one to have an understanding of the interactions between the lifestyle parameters in the data and in view of the demography identified in the survey.

However, for this study, my main objective will be to observe which demography has a better or improved work life balance with the available parameters in the data set.

Data Source and Description

The data set was made available by Yvon Dalat on kaggle.com. It consists of 15,977 survey responses with 24 attributes describing lifestyle and behavior of respondents. The latest version of the data was uploaded 14 March, 2021.

Limitations of the data

The data was collected through a survey of the respondents, and it's subjective of the respondents and as such it would be difficult to determine the veracity and their responses to the survey questions.

Steps followed

I imported the data set into Bigquery and which was used to wrangle, manipulate and prepare the data set for analysis. The resulting tables from Bigquery operations were then download onto my local computer and then imported into Power BI desktop, which I used to further transform the data particularly the floats were rounded to 2 decimal places before the data was loaded. I then used Power BI to visualize/analyze and summarize the output.

This exploratory analysis of the data was perform in view of 5 main context as below:

1. Healthy body: reflecting your fitness and healthy habits.
2. Healthy mind: indicating how well you embrace positive emotions.

3. Expertise: measuring the ability to grow your expertise and achieve something unique.
4. Connection: assessing the strength of your social network and your inclination to discover the world.
5. Meaning: evaluating your compassion, generosity and satisfaction living the life of your dream. We will consider meaning under passion and work life balance index.

It's important to mention that I got some insights on Power BI from @Kenji Explains and @Alex The Analyst. They can both be checked out on their youtube channels.

CLEANING THE DATA

Checking for the number of rows in the data set

```
SELECT COUNT(*)  
FROM Lifestyle_and_wellbeing.work_life_balance
```

Rows 15,972

Checking for NULL values in the data set

```
SELECT *  
FROM Lifestyle_and_wellbeing.work_life_balance  
WHERE FRUITS_VEGGIES IS NULL OR DAILY_STRESS IS NULL  
OR PLACES_VISITED IS NULL OR CORE_CIRCLE IS NULL  
OR SUPPORTING_OTHERS IS NULL OR SOCIAL_NETWORK IS NULL  
OR ACHIEVEMENT IS NULL OR DONATION IS NULL  
OR BMI_RANGE IS NULL OR TODO_COMPLETED IS NULL  
OR FLOW IS NULL OR DAILY_STEPS IS NULL  
OR LIVE_VISION IS NULL OR SLEEP_HOURS IS NULL  
OR LOST_VACATION IS NULL OR DAILY_SHOUTING IS NULL  
OR SUFFICIENT_INCOME IS NULL OR PERSONAL_AWARDS IS NULL  
OR TIME_FOR_PASSION IS NULL OR WEEKLY_MEDITATION IS NULL  
OR AGE IS NULL OR GENDER IS NULL  
OR WORK_LIFE_BALANCE_SCORE IS NULL;
```

No output given after running the above query, thus no NULL value present in the data set.

Checking for the types of values the columns in the data are taking

```
SELECT DISTINCT AGE, FRUITS_VEGGIES, DAILY_STRESS,  
                PLACES_VISITED, CORE_CIRCLE, SUPPORTING_OTHERS,  
                SOCIAL_NETWORK, ACHIEVEMENT, DONATION, BMI_RANGE,  
                TODO_COMPLETED, FLOW, DAILY_STEPS, LIVE_VISION, SLEEP_HOURS,  
                LOST_VACATION, DAILY_SHOUTING, SUFFICIENT_INCOME,  
                PERSONAL_AWARDS, TIME_FOR_PASSION, WEEKLY_MEDITATION,  
                GENDER, WORK_LIFE_BALANCE_SCORE  
FROM Lifestyle_and_wellbeing.work_life_balance
```

Manipulating the attributes of the data in line with the 5 context for this analysis

Healthy body

```
SELECT AGE,  
       GENDER,  
       BMI_RANGE,  
       Avg(FRUITS_VEGGIES) AS FRUITS_VEGGIES_SERVING,  
       Avg(DAILY_STEPS) AS STEPS_PERDAY,  
       Avg(SLEEP_HOURS) AS HOURS_SLEEP,  
FROM Lifestyle_and_wellbeing.work_life_balance  
GROUP BY AGE,  
       BMI_RANGE,  
       GENDER;
```

Healthy mind

```
SELECT AGE,  
       GENDER,  
       SUFFICIENT_INCOME,  
       Avg(DAILY_STRESS) AS STRESS_PERDAY,  
       Avg(FLOW) AS WORK_FLOW,  
       Avg(WEEKLY_MEDITATION) AS MEDITATIONS_PERWEEK,  
       Avg(DAILY_SHOUTING) AS DAILY_SULKING,  
FROM Lifestyle_and_wellbeing.work_life_balance  
GROUP BY AGE,  
       GENDER,  
       SUFFICIENT_INCOME;
```

Expertise

```
SELECT AGE,  
       GENDER,  
       Avg(PERSONAL_AWARDS) AS PERSONAL_RECOGNITION,  
       Avg(FLOW) AS WORK_FLOW,  
       Avg(ACHIEVEMENT) AS ACCOMPLISHMENT,  
       Avg(TODO_COMPLETED) AS COMPLETED_WORK,  
       Avg(DONATION) AS DONATIONS  
FROM Lifestyle_and_wellbeing.work_life_balance  
GROUP BY GENDER,  
       AGE;
```

Connection

```
SELECT AGE,  
       GENDER,  
       Avg(PLACES_VISITED) AS PLACES_BEEN,  
       Avg(CORE_CIRCLE) AS FAMILY_FRIENDS,  
       Avg(SOCIAL_NETWORK) AS CONNECTIONS,  
       Avg(LOST_VACATION) AS MISSED_HOLIDAY,  
       Avg(DAILY_STRESS) AS STRESS_PERDAY  
FROM Lifestyle_and_wellbeing.work_life_balance  
GROUP BY AGE,  
       GENDER;
```

Passion

```
SELECT AGE,  
       GENDER,  
       Avg(TIME_FOR_PASSION) AS HOURS_4_PASSION,  
       Avg(PERSONAL_AWARDS) AS PERSONAL_RECOGNITION,  
       Avg(LIVE_VISION) AS VISIONS,  
       Avg(FLOW) AS WORK_FLOW,  
       Avg(TODO_COMPLETED) AS COMPLETED_WORK  
FROM Lifestyle_and_wellbeing.work_life_balance  
GROUP BY AGE,  
       GENDER;
```

Work life balance index

```
SELECT AGE,  
       GENDER,  
       Avg(WORK_LIFE_BALANCE_SCORE) AS WORK_LIFE_INDEX  
FROM Lifestyle_and_wellbeing.work_life_balance  
GROUP BY AGE,  
       GENDER  
ORDER BY Avg(WORK_LIFE_BALANCE_SCORE) DESC;
```

Analysis/Visualization using Power BI

Summary of charts in Healthy body Dashboard

The line graph Average of STEPS_PERDAY by Age and Gender

The Less than 20 years old Males had up to 13.21% of Average STEPS_PERDAY.

Average of STEPS_PERDAY was higher for Male (5.88) than Female (5.61). Average of STEPS_PERDAY for Males and Females diverged the most where AGE was 36 to 50, when Males had average of 0.33 STEPS_PERDAY higher than Females.

Stacked column chart Average HOURS_SLEEP by Age and Gender

Less than 20 years old Females had up to 12.91% of Average HOURS_SLEEP.

Average of HOURS_SLEEP was higher for Females (7.10) than Males (6.91). Average of HOURS_SLEEP for Females and Males changed the most where AGE was Less than 20, when Females had an average 0.30 HOURS_SLEEP higher than Male.

Key Influencers charts: Average STEPS_PERDAY on BMI and Average HOURS_SLEEP on BMI

In this study BMI is given a range of 1 and 2; 1 = BMI < 25 while 2 = BMI > 25

When the Average of STEPS_PERDAY decreases the BMI range increases. Inversely, when the Average of STEPS_PERDAY increases the BMI range decreases.

Also when Average of HOURS_SLEEP decreases the BMI range increases, and when Average of HOURS_SLEEP increases the BMI range decreases.

The Decomposition tree chart: Average FRUITS_VEGGIES_SERVINGS on BMI

It is observed that among the demographic group in the survey, that those who had more average servings of fruits/veggies had a BMI of 1(which is a better BMI of the 2 rankings) while those who had less average servings of fruits/veggies had a BMI of 2.

Summary of charts in Healthy mind Dashboard

Line graph Average of MEDITATION_PERWEEK by Age and Gender

The 51 or more years old Males observed 13.97% of Average

MEDITATIONS_PERWEEK. Average of MEDITATIONS_PERWEEK was higher for Males (6.46) than Females (5.95). Average of MEDITATIONS_PERWEEK for Males and Females deviated most where AGE was between 21 to 35 years, when Males observed an average of 0.80 higher than Females in MEDITATIONS_PERWEEK.

Stacked column chart Average STRESS_PERDAY by age and Gender

The 21 to 35 year old group Males experienced up to 39.00% of Average

STRESS_PERDAY. Average of STRESS_PERDAY was higher for Males (5.21) than Females (3.04). Average of STRESS_PERDAY for Males and Females diverged the most where AGE was 21 to 35, when Males experienced an average of 9.82 STRESS_PERDAY higher than Females.

Clustered column chart Average SULKING_PERDAY by Age and Gender

Less than 20 years old group Females had up to 14.98% of Average DAILY_SULKING.

Average of DAILY_SULKING was higher for Females (3.19) than Males (2.74).

Average of DAILY_SULKING for Females and Males deviated the most where AGE was Less than 20, when Females had an average of 0.84 higher than Males in SULKING_PERDAY.

The Decomposition Tree chart: Average WORK_FLOW on SUFFICIENT_INCOME
In this study, sufficient income is ranked as 2, and less sufficient income is ranked as 1.

From the chart, among the demographic groups in the survey, those who observed more average work flow had sufficient income while those who observed less work flow had less income.

Summary of charts in Expertise Dashboard

Area chart ACCOMPLISHMENT by Age and Gender

Females in the Less than 20 years old group attained up to 13.38% of Average ACCOMPLISHMENT. Average of ACCOMPLISHMENT was higher for Males (4.07) than Females (4.06). Average of ACCOMPLISHMENT for Females and Males diverged the most where AGE was Less than 20, when Females attained an average of 0.26 ACCOMPLISHMENT higher than Males.

Line chart Average of COMPLETED_WORK by Age and Gender

51 or more year old Females achieved 13.81% of Average COMPLETED_WORK. Average of COMPLETED_WORK was higher for Females (5.89) than Males (5.50). Average of COMPLETED_WORK for Females and Males diverged the most where the AGE was 36 to 50, when Females achieved average of 0.44 COMPLETED_WORK higher than Males.

Ribbon chart Average of DONATIONS by Age and Gender

51 or more year old Females gave up to 16.29% of Average DONATIONS. Average of DONATIONS was higher for Female (2.84) than Male (2.46). Average of DONATIONS for Females and Males diverged the most where the AGE was 36 to 50, when Females gave an average of 0.44 higher than Males in DONATIONS.

Stacked area chart Average WORK_FLOW by Age and Gender

Average of WORK_FLOW was higher for Females (3.25) than Males (3.19). Average of WORK_FLOW for Females and Males diverged the most where the AGE was 51 or more, when Female had average 0.16 WORK_FLOW higher than Males.

Summary of charts in Passion Dashboard

Stacked bar chart HOURS_4_PASSION by Age and Gender

Less than 20 year old Males had up to 13.58% of Average HOURS_4_PASSION. Average of HOURS_4_PASSION was higher for Males (3.45) than Females (3.27). Average of HOURS_4_PASSION for Males and Females diverged the most where the AGE was Less than 20, when Males had an average of 0.52 HOURS_4_PASSION higher than Females.

Line graph Average of VISION by Age and Gender

51 or more year old Males had up to 14.24% of Average of VISIONS. Average of VISIONS was higher for Males (3.99) than Females (3.73). Average of VISIONS for Males and Females deviated the most where the AGE was 51 or more, when Males average of 0.33 higher than Females in terms of their VISIONS.

Summary of charts in Connections Dashboard

Ribbon chart Average of CONNECTIONS by Age and Gender

Female in the Less than 20 year old have up to 14.23% of Average CONNECTIONS. Average of CONNECTIONS was higher for Females (6.71) than Males (6.47). Average CONNECTIONS for Females and Males diverged the most where AGE was 21 to 35, when Females have 0.45 Average connections higher than Males.

Line chart Average of FAMILY_FRIENDS by Age and Gender

51 or more year old Females have up to 13.25% of Average of FAMILY_FRIENDS. Average of FAMILY_FRIENDS was higher for Females (5.74) than Males (5.20). Average of FAMILY_FRIENDS for Females and Males diverged the most where the AGE was 51 or more, when Females have an average of 0.74 higher than Males.

Stacked column chart Average MISSED_HOLIDAY by Age and Gender

Less than 20 year old Females had 14.16% of Average MISSED_HOLIDAY. Average of MISSED_HOLIDAY was higher for Males (2.98) than Females (2.87). Average of MISSED_HOLIDAY for Males and Females diverged the most where the AGE was 51 or more, when Males had 0.52 average MISSED_HOLIDAYS higher than Females.

Area chart Average of PLACES_BEEN by Age and Gender

21 to 35 year old Females have up to 13.51% of Average PLACES_BEEN. Average of PLACES_BEEN was higher for Females (5.30) than Males (4.99). Average of PLACES_BEEN for Females and Males diverged the most where the AGE was Less than 20, where Females had average PLACES_BEEN 0.46 higher than Males.

Summary of Work life balance index Visualization

Stacked bar chart Average of WORK_LIFE_INDEX by Age and Gender

51 or more year old group Females experienced up to 12.71% of Average WORK_LIFE_INDEX. Average of WORK_LIFE_INDEX was higher for Females (668.39) than Males (666.06). Average of WORK_LIFE_INDEX for Females and Males diverged the most where the AGE was 36 to 50, when Females experienced WORK_LIFE_INDEX 4.83 higher than Males.

CONCLUSION

The main observations on each of the five context areas for this study are as follows:

1. BMI is mostly impacted by quality of nutrition (fruits/Veggies) and physical activity (daily steps), and also has a correlation to the kind of quality of sleep

2. The major determinants of stress levels are the ability to focus on work flow, meditation and the sufficiency of income.
3. Those who achieve the most remarkable things, have also maximized their ability to complete daily todo list, focus on work flow and have earn a number of personal awards and recognition.
4. Having a core circle of family, close friends and networks contributes to the number of new places people visit, reduces daily stress and improve social connections.
5. People find more time for passion when they are able to complete well their daily todo list, work flow through the day and have obtained many personal awards and recognition.
6. From the observation of the work life balance score(index), women appear to have better work life balance than men.