

HAYEK GLOBAL COLLEGE

Data Driven Analysis – Final Report

Professor Igor Agatti Lima

Group Assignment: The CEO enjoyed the presentations and loved the work done. In one of his fit of crazy/genius decisions he decided that North Carolina is not the state that the new business must be implemented. He interpreted one of his dreams and decided to change it to South Carolina. He also implemented a new policy where team leaders will rotate groups, and the business ideas will rotate with the team leaders.

The groups are now given the option: **(i)** Recreate the analysis report and presentation using the data from BRFSS. Although, instead of analyzing North Carolina the group will analyze South Carolina. **(ii)** Review the reports and presentation previously made on North Carolina, seeking new insights and improving on the work that was previously done. Each task force will prepare a document with relevant graphs and tables and give the description and written analysis of each graph and table.

For this deliverable, you should: **(i)** Create a final report; **(ii)** Create a group presentation that lasts no longer than 15 minutes. The report may be turned in after class, but the presentation will be during class hours. Everyone in the group is expected to present. If you cannot make it to class you can send a video.

Main project team: David Maycon, Felipe Pessoa, Pet'ka Špaková and Timothy Morenikeji

Review team: Anikó Barcziová, Elena Skalovskaia, Felipe Pessoa and Gabriela Becheanu

REPORT

Summary

1. According with what was discussed in Data Driven Analysis course, the students were given an .xlsx document containing the data gathered for the Behavioral Risk Factor Surveillance System (BRFSS) for the year of 2020 – published in July 7th, 2021. The BRFSS is a collaborative project between all of the states in the United States and participating US territories and the Centers for Disease Control and Prevention (CDC), as described in its overview, also given for the students for the purpose of this report.

2. The case which was given described the students were just hired by Global ABC Company, and the CEO discovered BRFSS' publicly available data and assigned them to be part of a taskforce to find new business opportunities taking into consideration what was found for the state of North Carolina. There was no specification on the nature of the business, but only the task of analyzing the data and coming up with a business idea, even if such idea already existed.

3. After the initial presentation, there was a group change based on the “group assignment” reported above and the need for a final report, which was said it could be consolidated in only one document. In the next lines, the data analyzed and business idea will be described.

Data gathered from the BRFSS (2020) for North Carolina

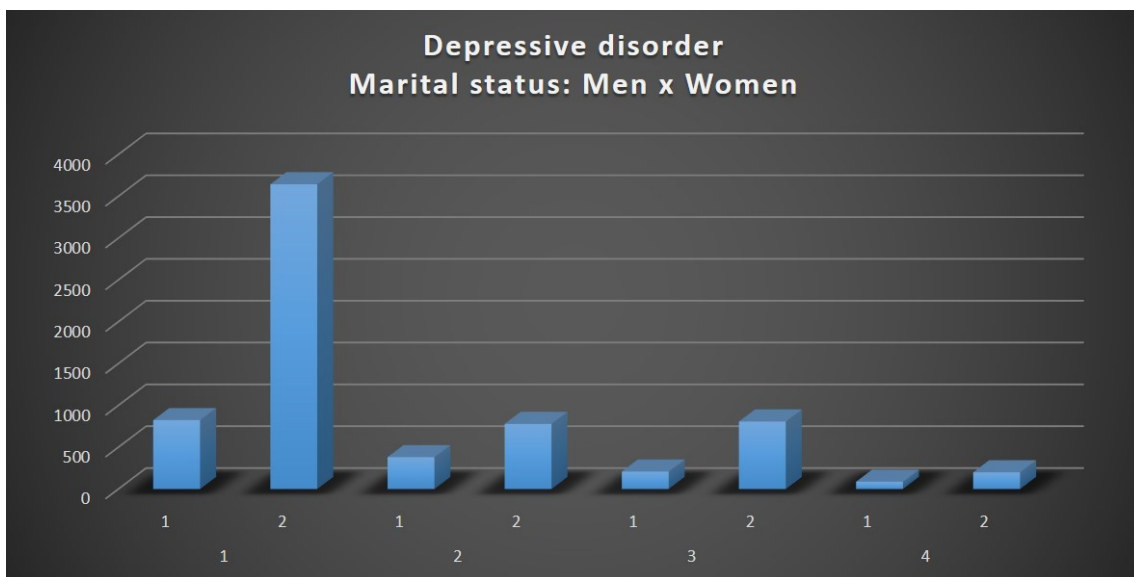
4. Before we can start the description of data found for North Carolina, we must warn that every graph used in this report is used also in the group presentation sent as annex of the groups' final submission. Also, as a final and consolidated document, we shall cover in only one section both the first gathering of data and the review made by the second team; always, of course, indicating where the points reviewed, challenged and the conclusions that could be taken. For any better schematization of such data, we encourage the reader to have a look at the annex to this report.

5. The main project team, in the stage of extracting data, had absolutely no previous hypothesis concerning what should be searched or which business would be interesting for the project. Therefore, the business idea proposed came entirely from data which called the

attention of the team members precisely during the extraction stage to a certain level of randomness.

6. Starting with the information that provoked the team to further inquiry, see **Graph 1** below, on the assessment of depressive disorder in men and women filtered by the marital status. For the image shown, the first-base numbers¹ are 1 (men) and 2 (women), while the second-base² are 1 (married), 2 (divorced), 3 (widowed) and 4 (separated).

(Graph 1)



7. As we can see, there is a noticeable disproportion on the occurrence of depressive disorder in married women. All the others, although women are still leading in depression cases, are fairly distributed in comparison the men in the divorced, widowed and separated groups.

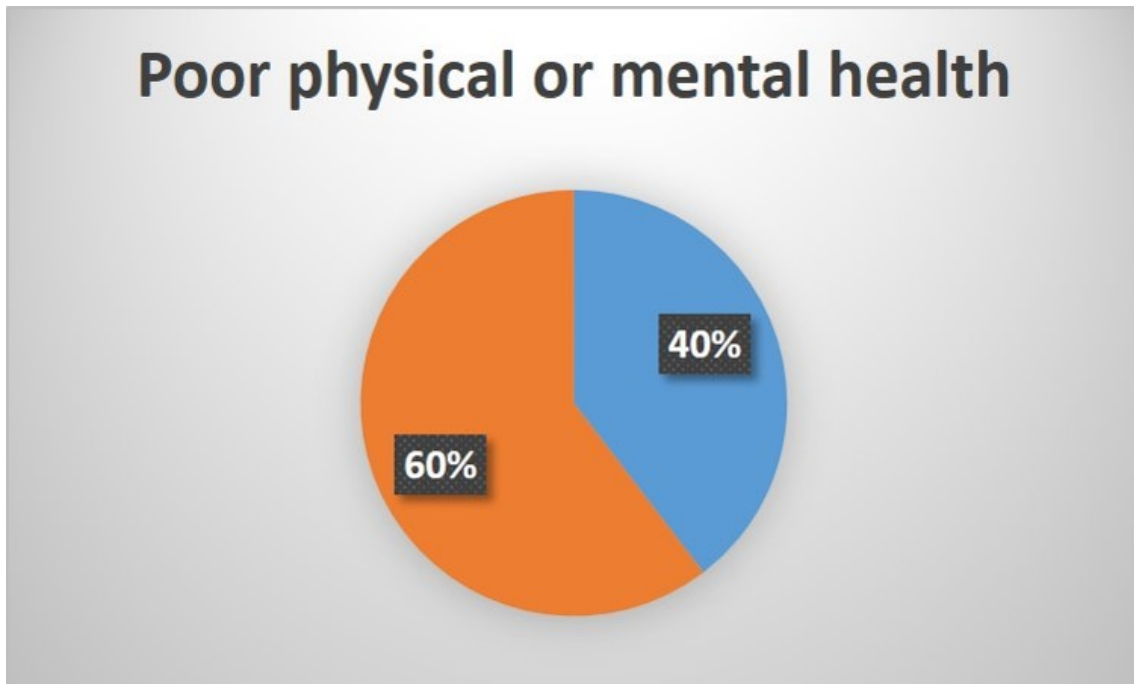
8. The main project team also noticed that women's numbers being higher than men's might be related to the fact that the BRFSS (2020) data had more responses from women (54.24%) than from men (45.76%), as shown in the Codebook Report. Although that probably explains the general women's numbers in comparison to men, the huge disproportion observed in the married women group is too deviant from the normal variation in order to be just a consequence from the difference of the amount of respondents from each sex.

9. And the disproportion of the identified group could be verified after analyzing other data, both more general and more specific. For this, **Graph 2** below shows that a majority of women (orange) report having both poor physical and mental health, against 40% from men:

¹ Meaning, the ones which appear lined closer to the graph bars.

² Meaning, the ones which appear lined right below the first-based numbers.

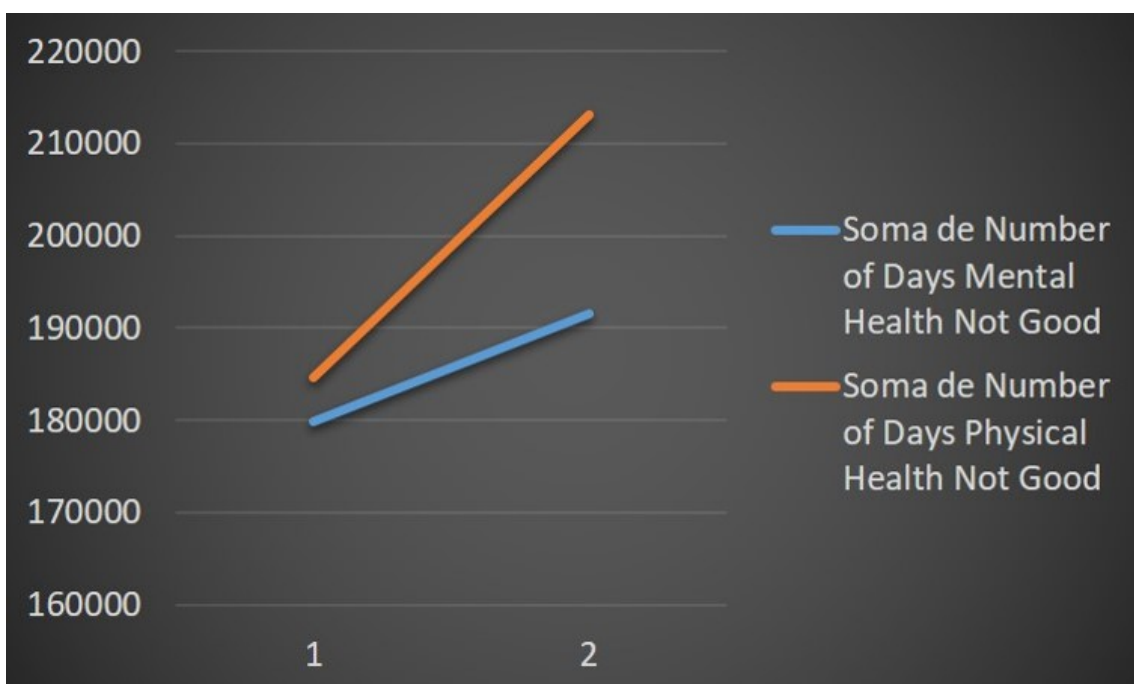
(Graph 2)



10. Such variation, although might give some hints on the data observed from depression in married women, falls into the variation of respondents. In other words, the majority of women reporting poor physical and mental health might not be the real indication that women, in general, are more prone to health problems.

11. The same conclusions can be taken from **Graph 3** below:

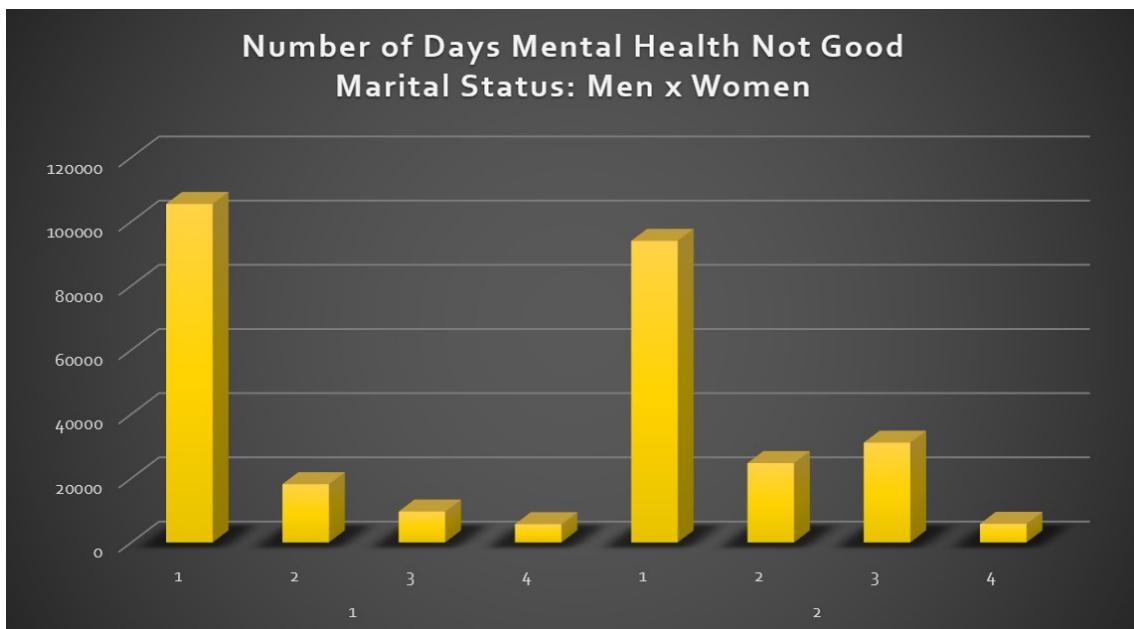
(Graph 3)



12. **Graph 3** shows a majority of women (orange) reporting a higher number of days which mental health was not good – against a smaller impact on men (blue). According to the Codebook Report, such data is related not only to depression, but also stress and other problems with emotions. Therefore, when we analyzed mental health issues of women in general, the distribution shows only the already expected variation from men’s group.

13. The same data can be seen in **Graph 4**, where we see the distribution of the number of days which mental health was not good in men and women, but also controlling for the marital status. Notice we see that the first-base numbers are 1 (married), 2 (divorced), 3 (widowed) and 4 (separated), while the second-base numbers are 1 (man) and 2 (women), all of them fairly distributed, only with a higher impact in married men and women.

(Graph 4)

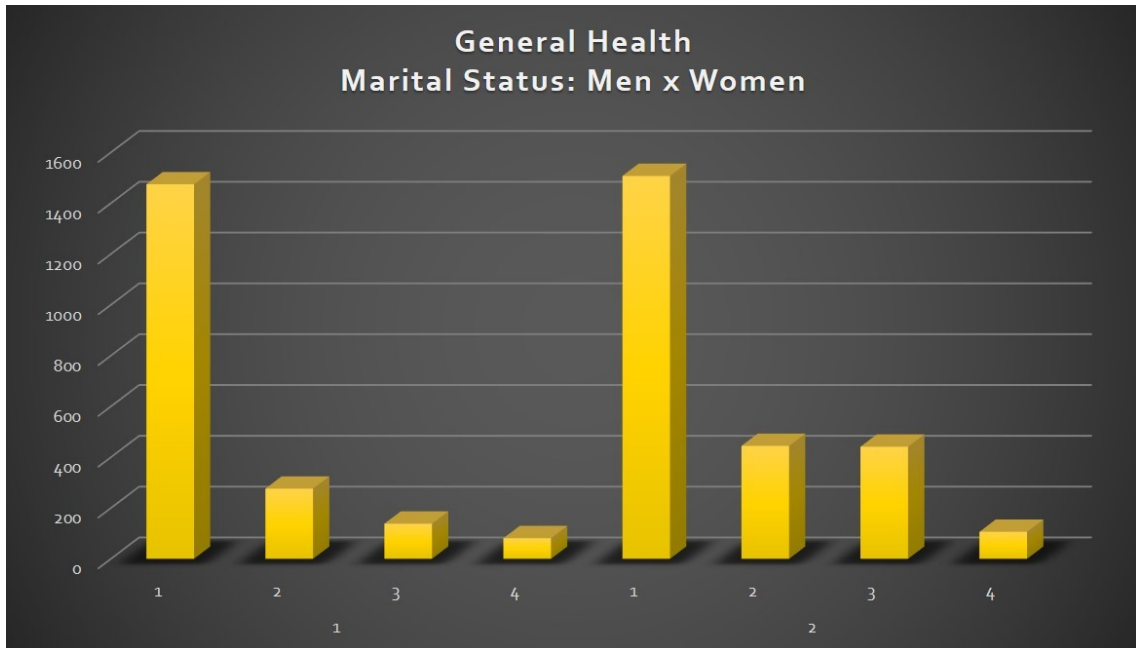


14. Thus, according to such data, at least there are no clear indications that women are more prone than men to have mental health issues. There is particularly no real or even noticeable correlation between female gender and the incidence of emotional disorders of every sort. The further the team analyzed, making the filters more and more broad, no other data showed that women in general might be more prone than men to any other health issues.

15. And such broader reasearch was conducted by taking the the data from respondents which were not exclusively mental health issues, but were evaluating general health. According to the Codebook Report, general is really a large concept, that can encompass basically any personal perception of health.

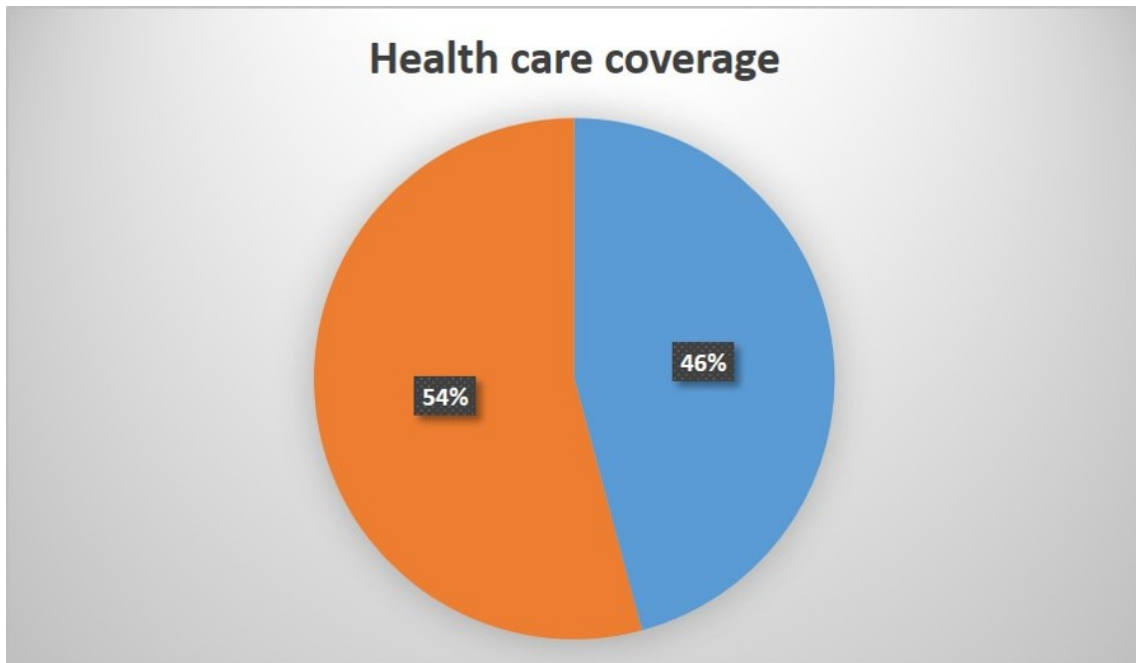
16. See **Graph 5** below on this matter:

(Graph 5)



17. In the same line, see **Graph 6**, which shows the variation between health care coverage of women (orange) and men (blue):

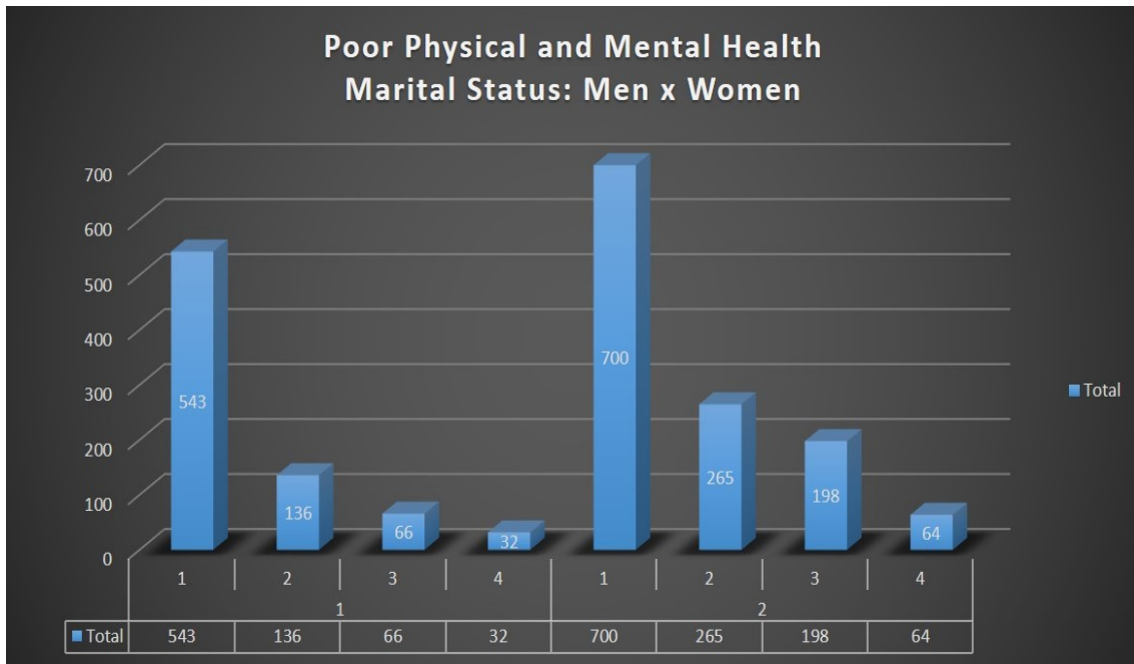
(Graph 6)



18. As seen before, the little variation might be satisfactorily explained by the variation of gender of the respondents.

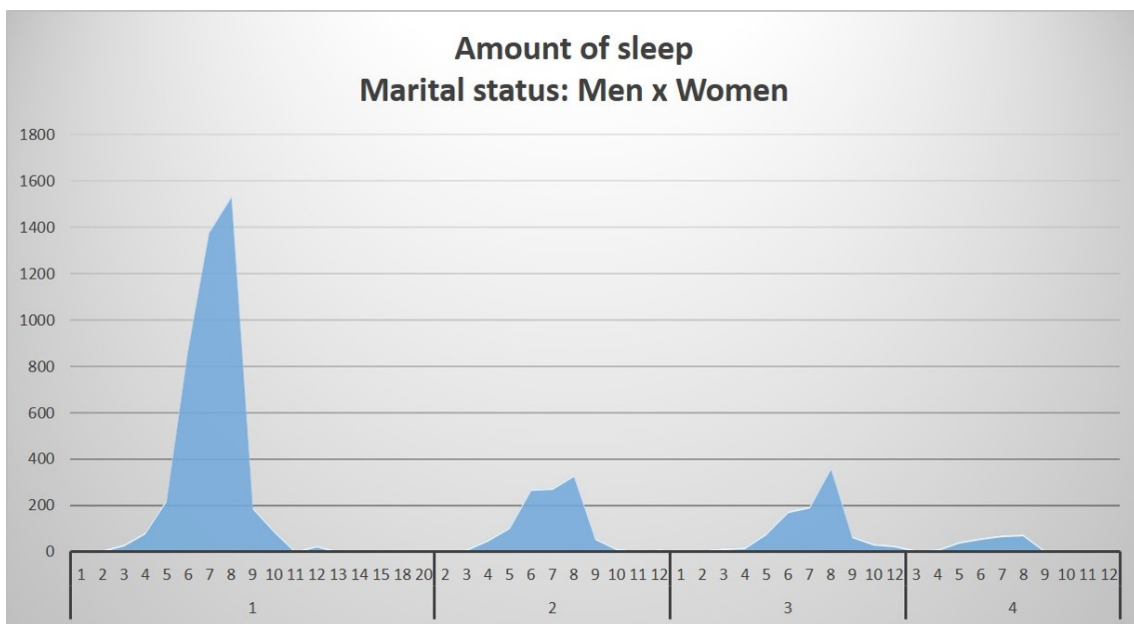
19. The same conclusion comes when we analyze **Graph 7**, which covers poor physical and mental health of men and women, just like Graph 2, but now controlling by marital status. For first-base numbers we have 1 (married), 2 (divorced), 3 (widowed) and 4 (separated), while the second-base numbers are 1 (man) and 2 (women)

(Graph 7)



20. We get no further information when we try to control the amount of sleep with marital status, which is also fairly distributed, as seen in **Graph 8**:

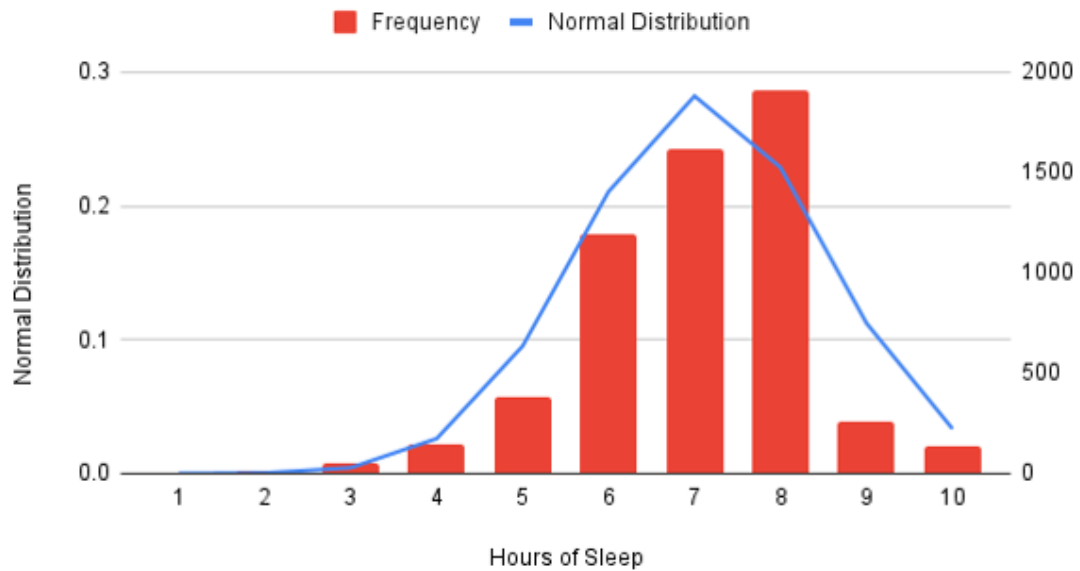
(Graph 8)



21. For the amount of hours of sleep, the main project team raised the hypothesis that a noticeable variation on sleep might indicate at least the clue of a symptom. Controlling for the frequency and normal distribution, no further information was found. For this, see **Graph 9**:

(Graph 9)

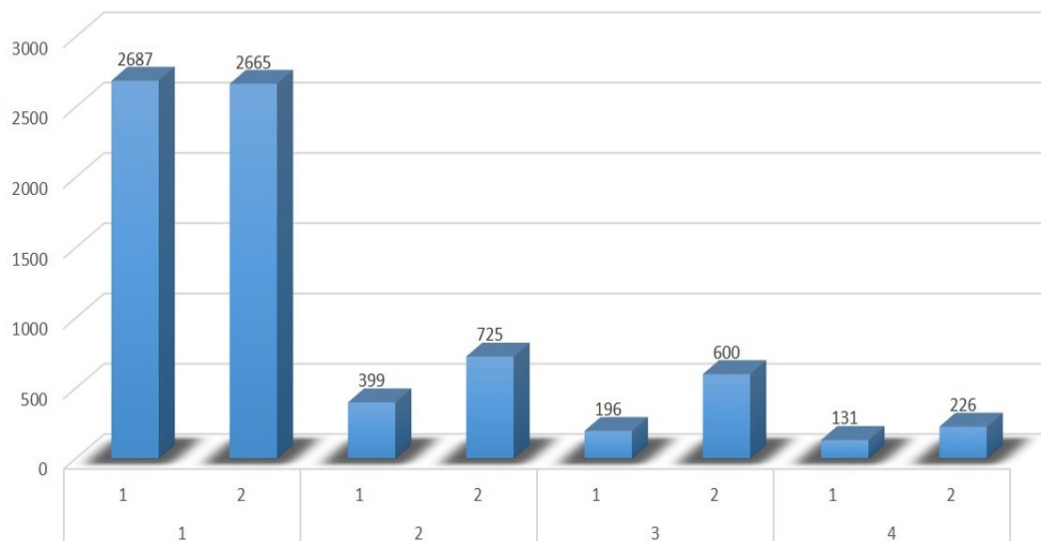
Normal Distribution vs. Hours of Sleep



22. Now, see **Graph 10**:

(Graph 10)

Children living in household Marital status: Men x Women



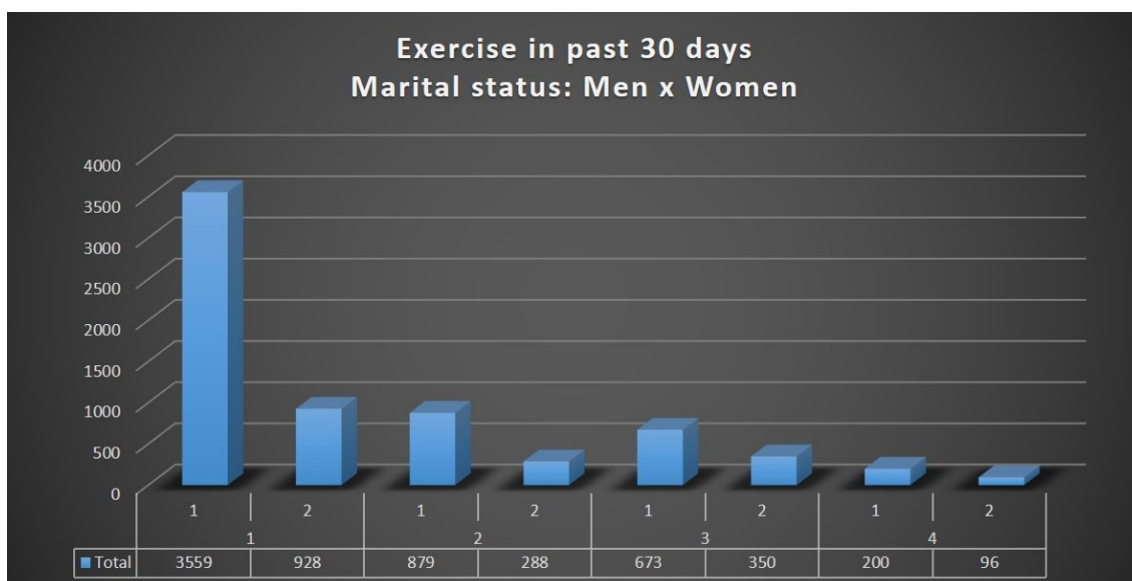
23. For **Graph 10**, the main project team was still trying to figure some pattern which could justify or at least throw some light into the data that was found in Graph 1. For this, the team raised the hypothesis that depression and other mental disorders might be related to children, and so the students tried to control for such group. Notice that first-base numbers are 1 (man) and 2 (women), while second-base are 1 (married), 2 (divorced), 3 (widowed) and 4 (separated).

24. As we see in **Graph 10**, there is fair distribution for both men and women in all groups, which might raise at least the curiosity that a fair proportion of children between men and women – specially in the married group – might be precisely related to the explanation for the disproportion of depression in women. Then again, this is just a hypothesis, or a line of observation, which might be raised, although not confirmed by the data gathered.

25. Another interesting data that was found trying to control for other groups is the comparison of physical exercises of men and women by marital status. As we can see in **Graph 11** below, women generally exercise less than men, but we can see a disproportion of less exercises of married women in comparison to married men. As we will make the claim further on, according to studies which extrapolate the data found in the boundaries of this project, such numbers might be related to the amount of free time women have for themselves in marriage, particularly when related to gender role expectations in wedlock.

26. For **Graph 11**, first-base numbers are 1 (man) and 2 (women), while second-base are 1 (married), 2 (divorced), 3 (widowed) and 4 (separated):

(Graph 11)



27. Making further research in order to explain the phenomenon identified in Graph 1, the main project team found scientific articles pointing out precisely what was seen. One of the lines of interpretations found were *“that women who indicate that a conflict in sex role expectations exists in their marriage report more symptoms of depression than married women who do not indicate such a discrepancy in sex role expectations.”*³ Another explanation, derived from a different studies, showed that females have larger and stronger social support networks than men, whereas men often report their wives to be their chief source of social support.⁴

28. After the first group presentation in class, the review team was assigned to challenge both the data and the assumptions made previously. The business hypothesis raised by the main project team were:

28.1. Factors related to married women’s depression might be **(1)** social network of support and/or **(2)** conflict in sex role expectations in wedlock.

28.2. A virtual application that **(1)** has social inputs and **(2)** trackings with healthcare specialists might be helpful to mitigate potential impacts of depression.

29. The review team, taking into account that a close review of every data reported so far showed exactly the same results, decided to control for other factors which might indicate that the married women group might be affected by other variables not seen yet.

30. For this, the review team started an deeper inquiry motivated by social and economic factors, formulating the following hypothesis:

30.1. Is this phenomenon related to education level?

30.2. Is this phenomenon related to income level?

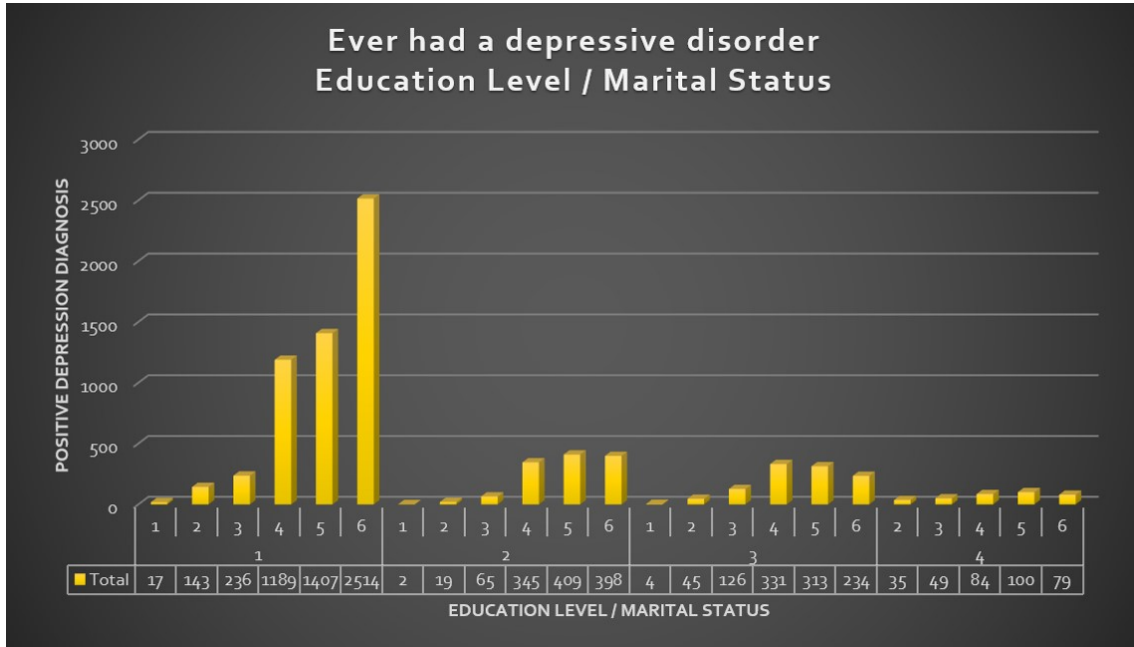
31. Starting with education, **Graph 12** below tries to control depressive disorder both for sex, marital status and education to assess if there are any discrepancies not observed before. We must notice that education level is shown in the image on first-base numbers, which are 1 (never attend school), 2 (elementary education), 3 (incomplete high school), 4 (high school

³ KRAUSE, Neal. “Married women and depressive symptoms”. **Sociological Focus**. Vol. 15, No. 1 (January, 1982), pp. 25-40 (16 pages). Available at: < <https://www.jstor.org/stable/20831224> >

⁴ CANTECH LETTER. “Marriage means less chance of depression for men, opposite for for women: study”. Available at: < <https://www.cantechletter.com/2017/07/marriage-means-less-chance-depression-men-opposite-women-study/#:~:text=posts%2C%20Health%2C%20Science-,Marriage%20means%20less%20chance%20of,men%2C%20opposite%20for%20women%3A%20study&text=On%20the%20mental%20health%20front,%2C%20widowed%2C%20separated%20or%20divorced.> >

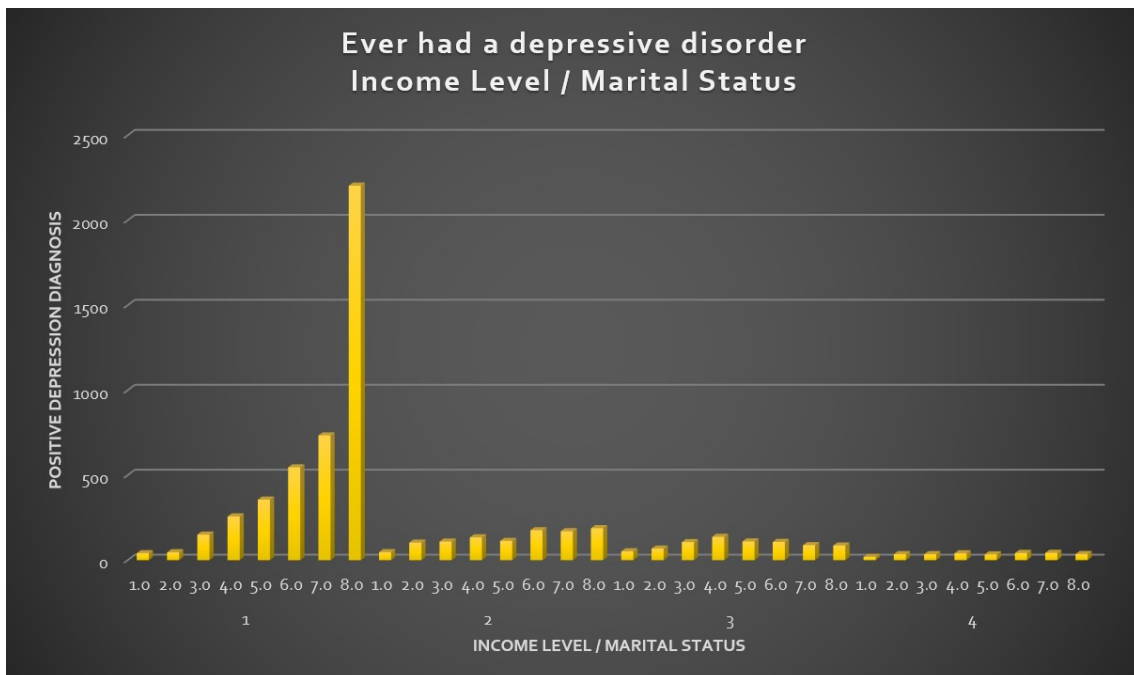
graduate), 5 (incomplete college) and 6 (college graduate); for second-base numbers we shall have 1 (married), 2 (divorced), 3 (widowed) and 4 (separated).

(Graph 12)



32. Following the second hypothesis, we have Graph 13:

(Graph 13)



33. For this last graph above, there was the control for income level. For first-base numbers we have 1.0 (makes less than \$10,000), 2.0 (makes more than \$10,000 and less than \$15,000),

3.0 (makes more than \$15,000 and less than \$20,000), 4.0 (makes more than \$20,000 and less than \$25,000), 5.0 (makes more than \$25,000 and less than \$35,000), 6.0 (makes more than \$35,000 and less than \$50,000), 7.0 (makes more than \$50,000 and less than \$75,000) and 8.0 (makes more than \$75,000), while second-base numbers are 1 (married), 2 (divorced), 3 (widowed) and 4 (separated).

34. It is very simple to observe that reporting of depression increases with both the increase of education and income level. However, both these data might not support neither conflict the data seen in Graph 1, particularly because they might solely indicate that people with higher education and more money might be more willing to look and pay for mental health treatment.

35. The conclusions that both main project and review teams reached are:

35.1. Married women are more prone to depression than any other controled group identified in the study.

35.2. No other data from the BRFSS (2020) helps supporting or contradicting the conclusion expressed in 35.1.

35.3. Further conclusions concerning the phenomenon seen in Graph 1 must extrapolate the boundaries of this study.

Business proposition

36. According to all the data gathered, and the interpretation given to it, the main project team – with the further inquiry of the review team – proposed a business concerned with depression and aimed mostly at married women, but not excluding other controled groups of women and neither men.

37. The business consists in a phone application, with both free and paid versions, that helps with depression with daily activities and social bonding, but also keeps track with physicians so that people diagnosed with depression might have an around-the-clock checking and reporting. The main goal is to lower to a minimum the tendency of depression by launching an app that will have features such as: **(1)** alarms and to-do lists in order to make the user turn from inactive to active; **(2)** a tracking of success of the user in performing his or her activities and shall record the inputs for a future report that might be integrated with physicians' or hospital's systems, so results can be cross-checked when in person consultations; **(3)** a channel for contact with personal doctors to make appointments.

38. Such business is already a reality in Brazil with the Conemo.⁵ The first prototype being developed by the University of São Paulo in partnership with a Peruvian university. Similarly to this business proposition, Conemo is an app that proposes activities that encourage certain behaviors and set goals, while the entire process is monitored by trained nurses. The first trials of the app have shown to be successful with patients reporting they manage to overcome depression with the solution proposed.

39. In this sense, the target customers of this business proposition shall be women between 30-45 years that are in their productive years and prefer both fast and easy advice. These women are also urban citizens, preferring digital innovations and wanting to solve their issues as quick as possible. As a secondary public, since the solution can be adapted to other groups, are men between 35-45 who are seeking for emotional support and prefer digitalisation.

40. With these solutions, and taking into consideration the data given, both the teams engaged in the project believe the proposed business can offer real and palpable good for the target customers.

⁵ CNN BRASIL. “Aplicativo desenvolvido pela USP ajuda no combate à depressão”. Available at: <
<https://www.cnnbrasil.com.br/saude/aplicativo-desenvolvido-pela-usp-ajuda-no-combate-a-depressao/>
>