

Diabetes treatment and possible factors affecting its effectiveness

Web Scraping Project

By
Oluwole Alowolodu

Outline



Introduction



Scrapy and Website



Methods



Data visualization



Discussion & Conclusion

Diabetes

8.5% prevalence in adult population

In US, (9.4%) 30.3 million people live with diabetes according to CDC

Avandia Oral

Metformin

Invokana

Method – WebMD \$ Scrapy

- WebMD is an online source for health-related information
- Suitable for this project because it lists medical conditions and drugs
- Huge user review database which includes;
 - Medical condition
 - Date
 - Duration of treatment
 - Age range of user
 - Gender
 - Ratings based on effectiveness, ease of use and satisfaction
 - Comments
- Scrapy, an open source web scraping framework was used to extract data of 3 diabetes drug.

Condition: Type 2 Diabetes Mellitus

6/30/2018 1:16:35 PM

Reviewer: 55-64 on Treatment for 1 to 6 months (Patient)

Effectiveness



Ease of Use



Satisfaction



Comment:

I started taking this twice a day. 500 mg twice a day. Before I started taking it my a1c was 7.3. After 3 or 4 months my a1c is now 6.2. I had no bad side effects at first. My bowel movements were hard but that's about it. It has now been 3 or 4 months and I am sick to my stomach. I have explosive diarrhea now and cannot leave the bathroom. I have to work for a living so I had to quit taking ... [Show Full Comment](#)

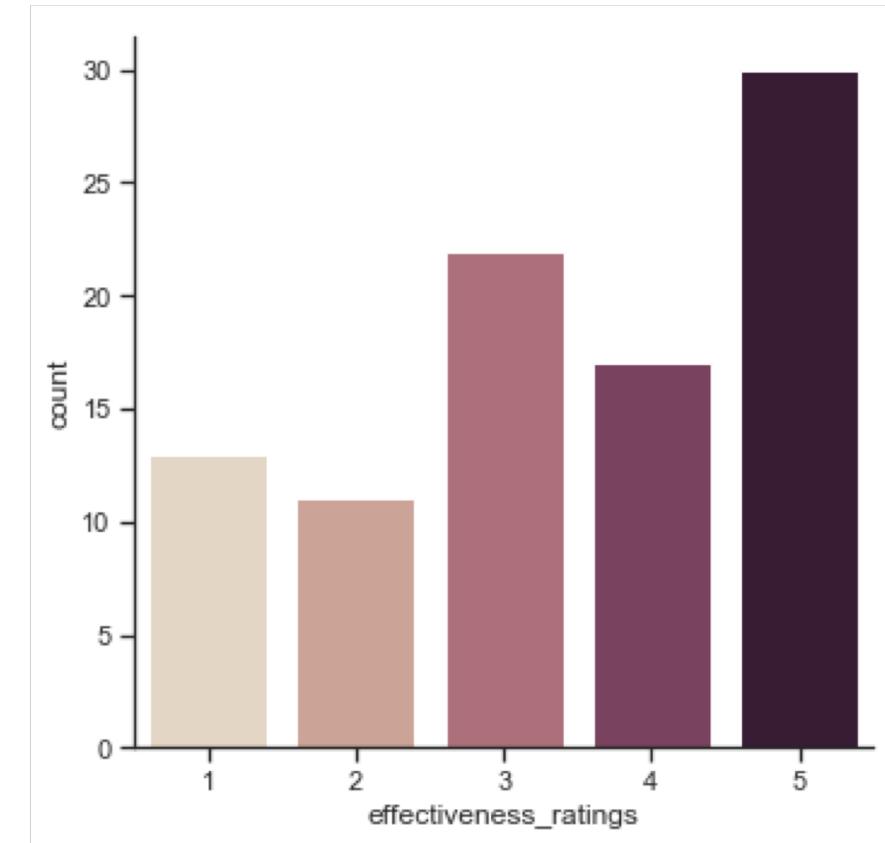
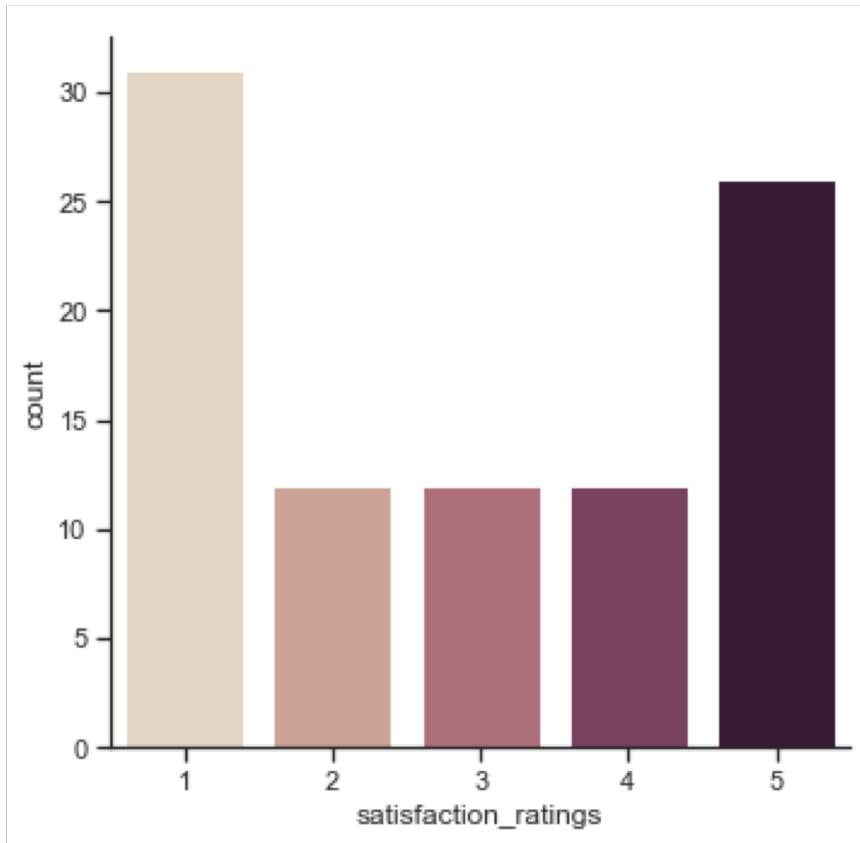
11 people found this review helpful.

Was this review helpful? [Yes](#) | [No](#)

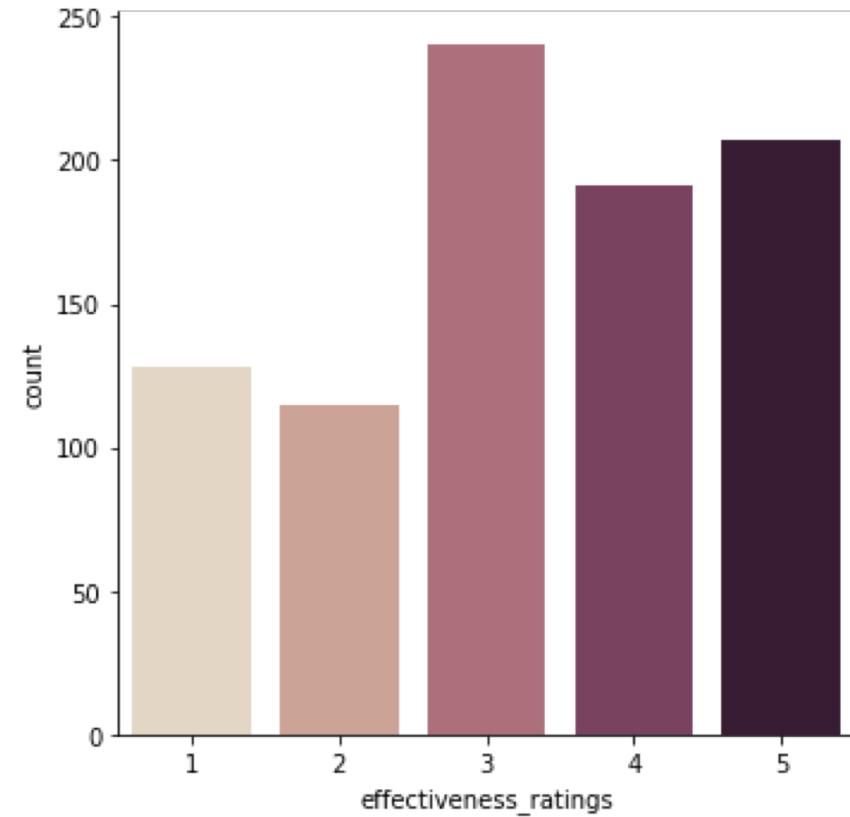
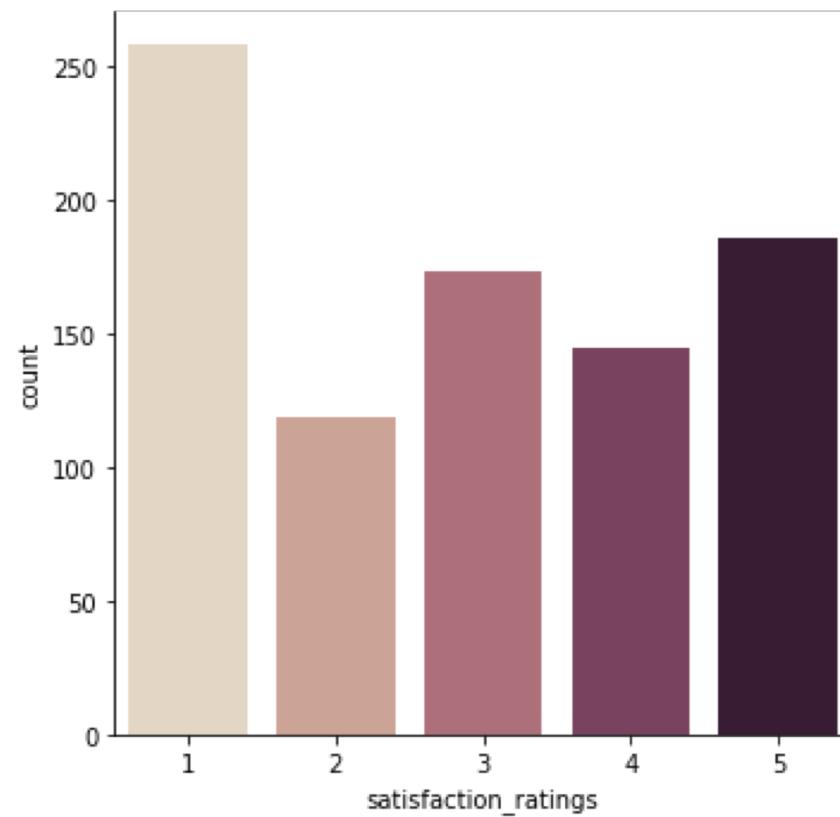


[Report This Post](#)

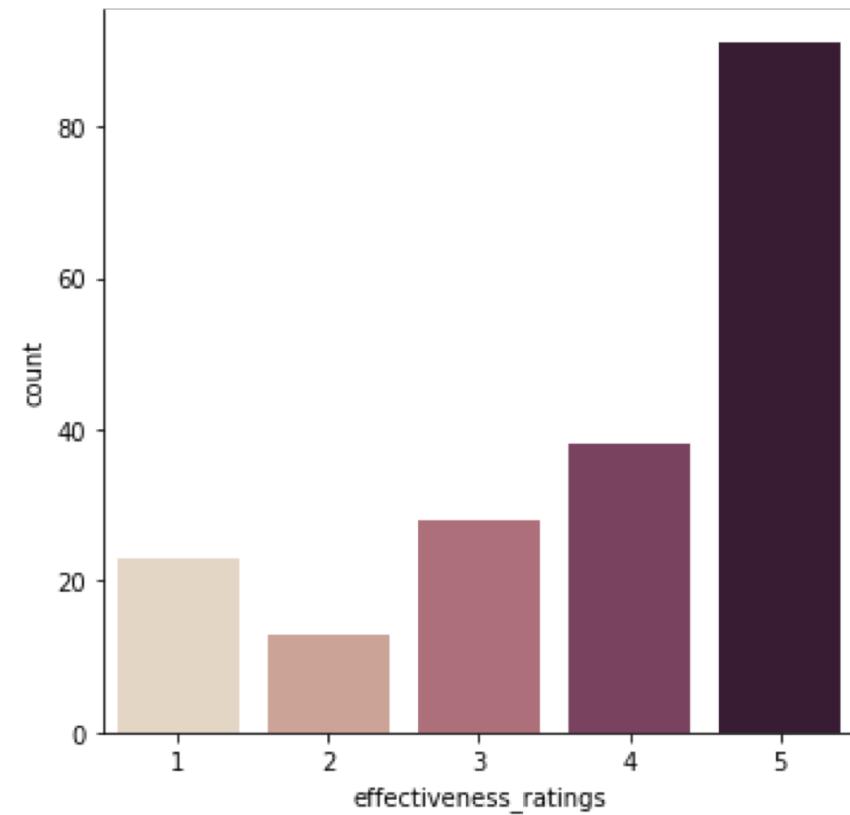
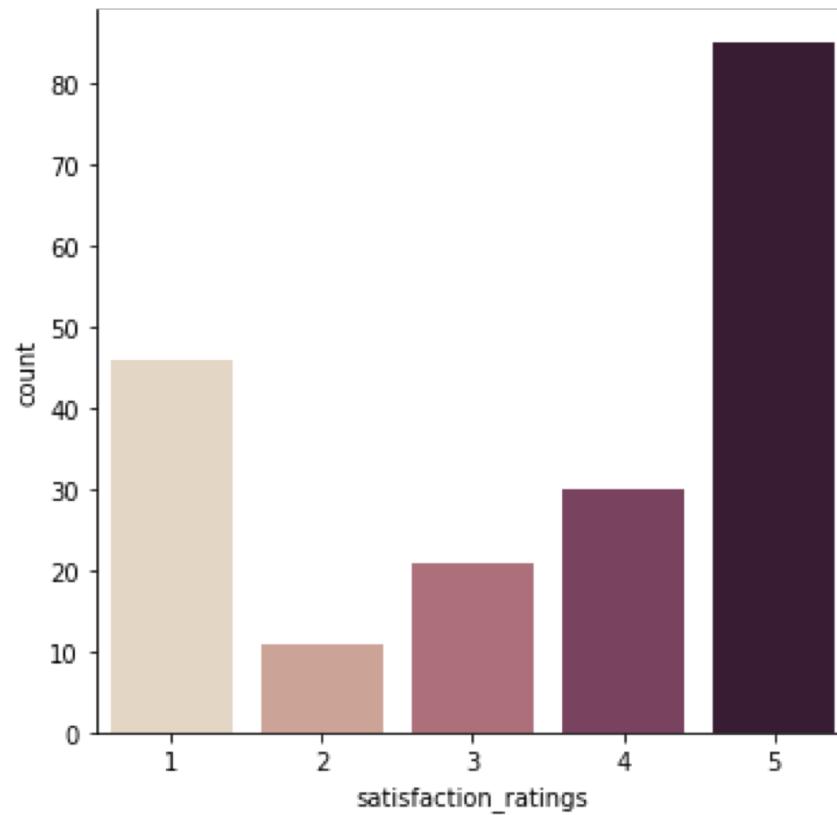
Avandia Oral



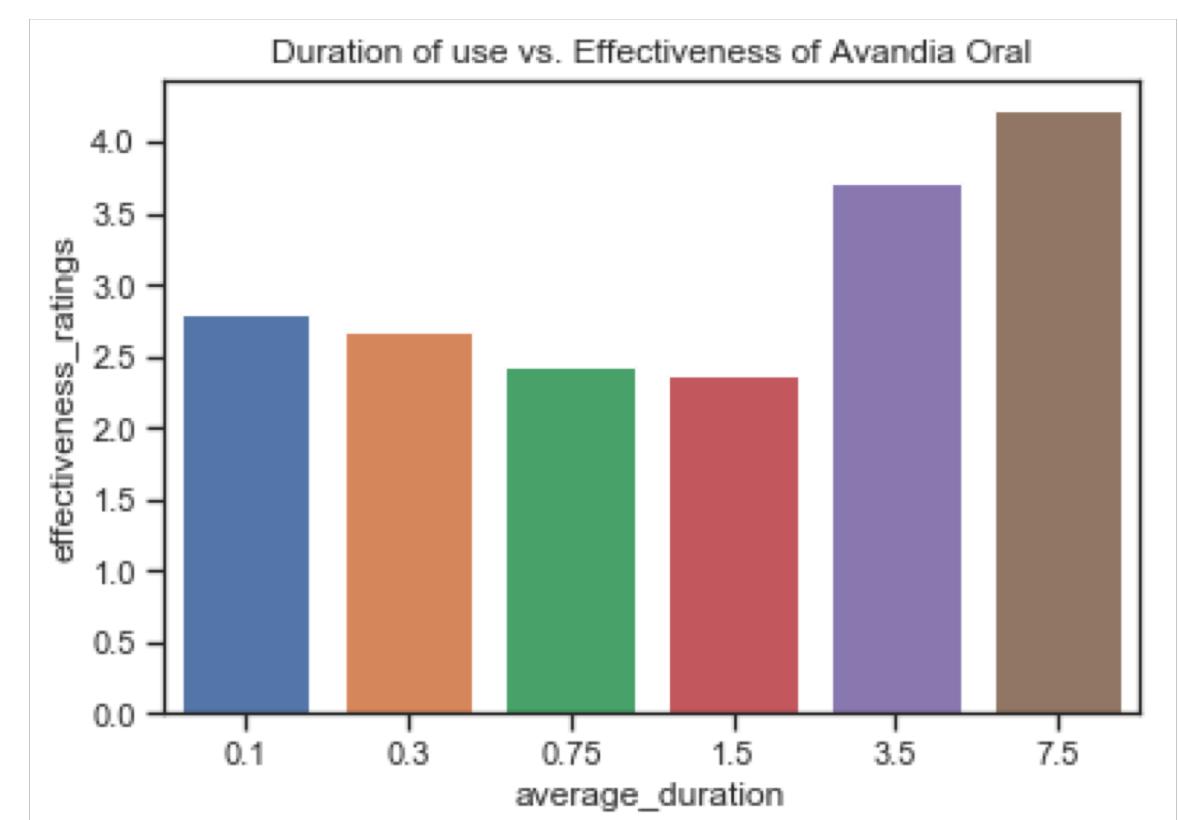
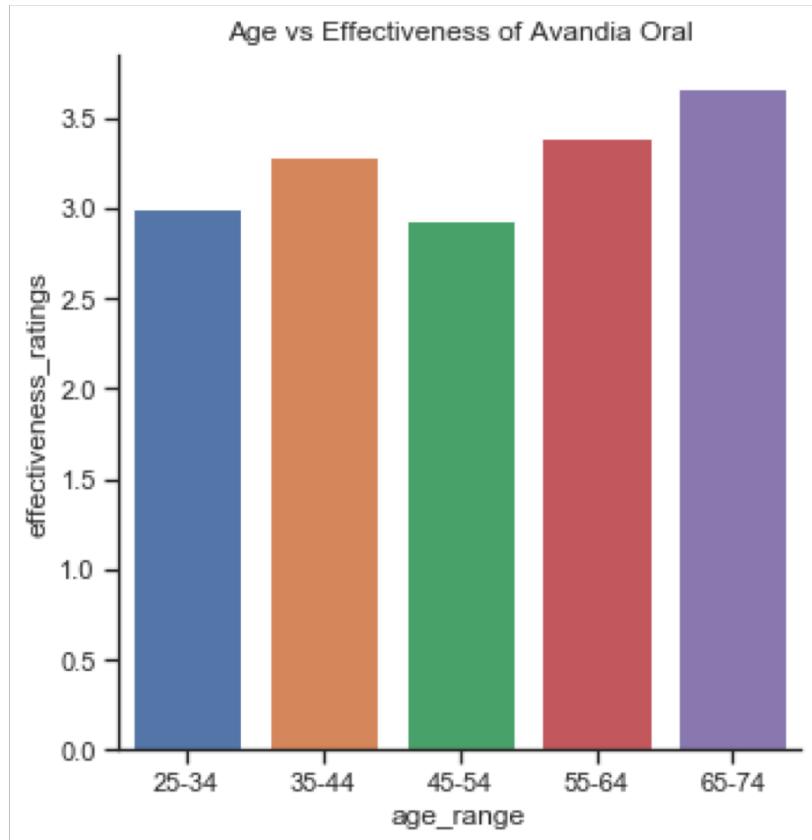
Metformin



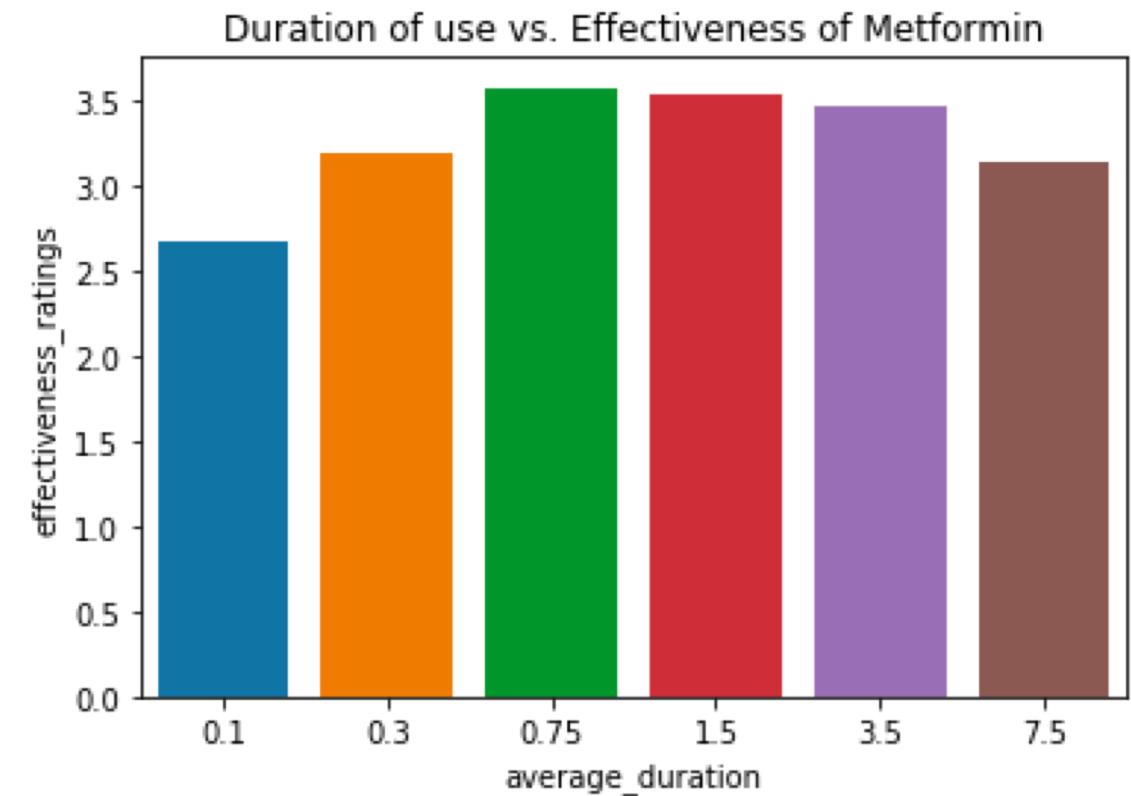
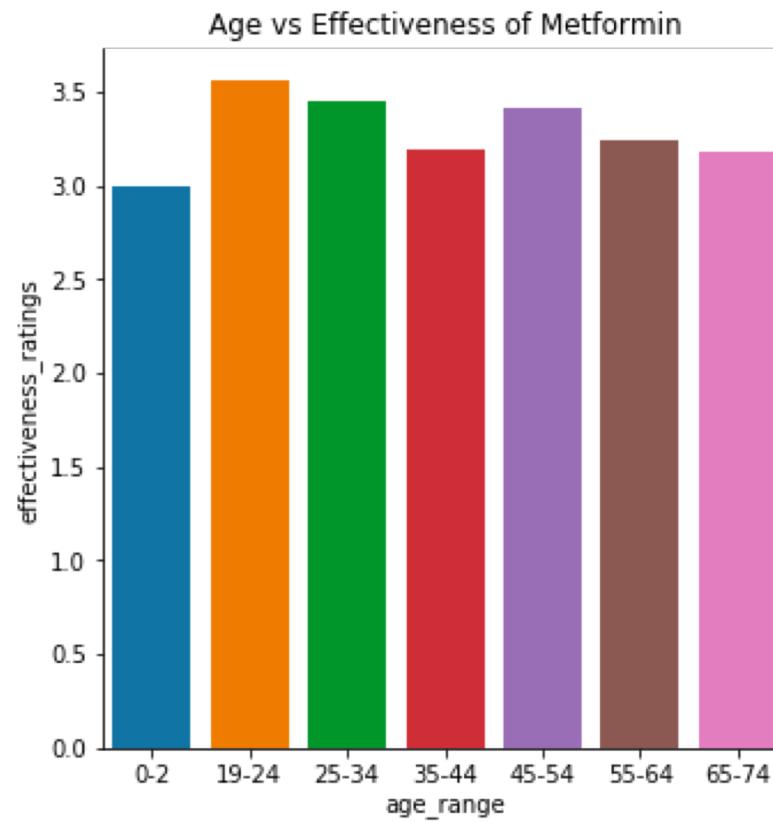
Invokana



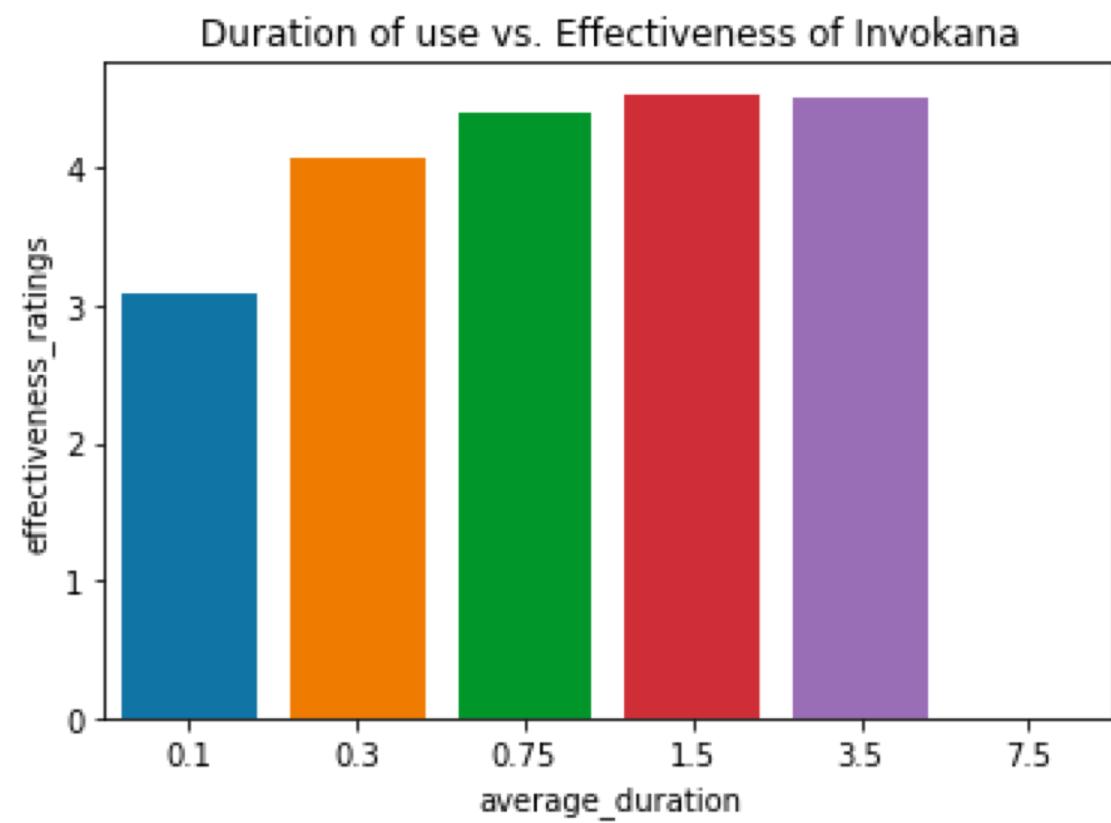
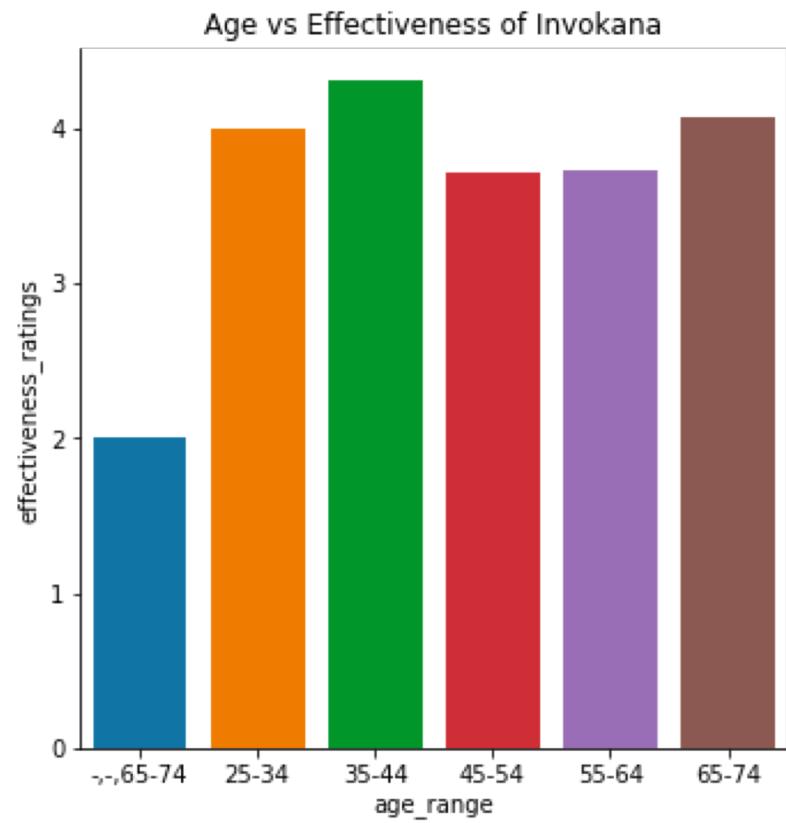
Visualization of possible factors –Avandia oral



Visualization of possible factors –Metformin

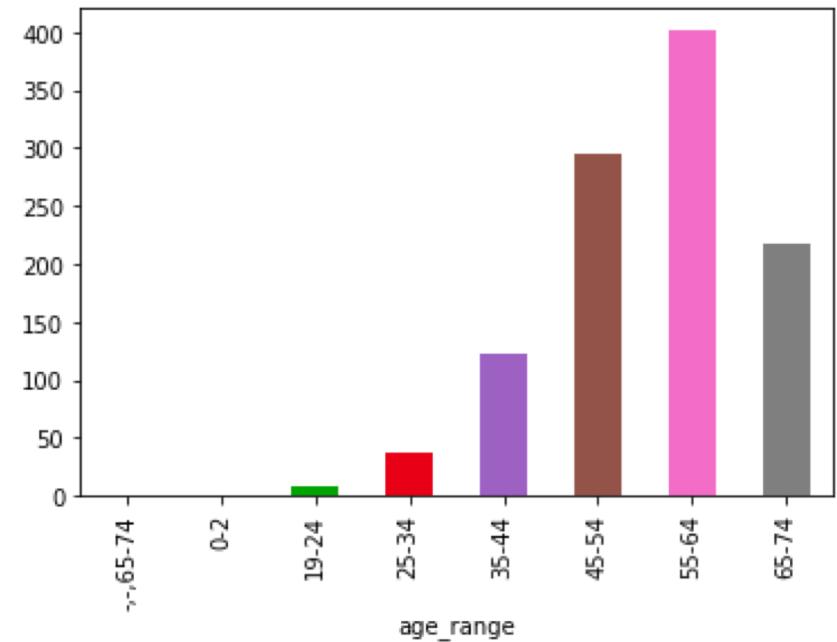
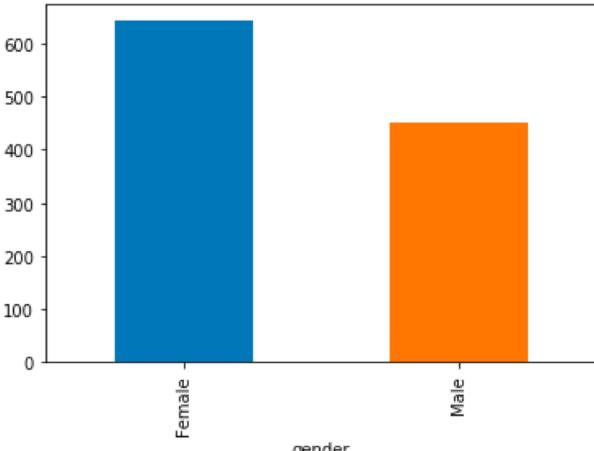


Visualization of possible factors - Invokana



Discussion

- As displayed in the plots, the data collected and analyzed indicates the demographic make-up of most diabetes patients are mostly between the ages of 45-74, and females.



Discussion



Various factors may have contributed to the ratings by the public, particularly the effectiveness ratings of the treatment, however, duration of treatment tends to be directly proportional to effectiveness.



Future analyses should include sentiment analyses of comments submitted to observe if there is a trend or repeated reason for non-continuation of treatment.