# HEALTHCARE EXPENDITURES FOR DIABETICS IN THE UNITED STATES

DOMINIC GARCIA

# AGENDA

Business Problem

Data

Visualizations & Insights

Conclusions

Future Work

# BUSINESS PROBLEM

- Beyond Type I is a nonprofit with a focus on education, advocacy & a path to cure Type I Diabetes.
  - A primary focus of their advocacy is the added cost of living for those diagnosed with the condition.
- Recently, they've aimed to put together a comprehensive report on the factors that determine how much income diabetic & non-diabetic citizens spend on healthcare. Specifically, they'd like to answer the following question:
  - **Are the demographic factors (race, sex, education level, income level, age) most important to determining money spent on healthcare the same for diabetics & non-diabetics?**

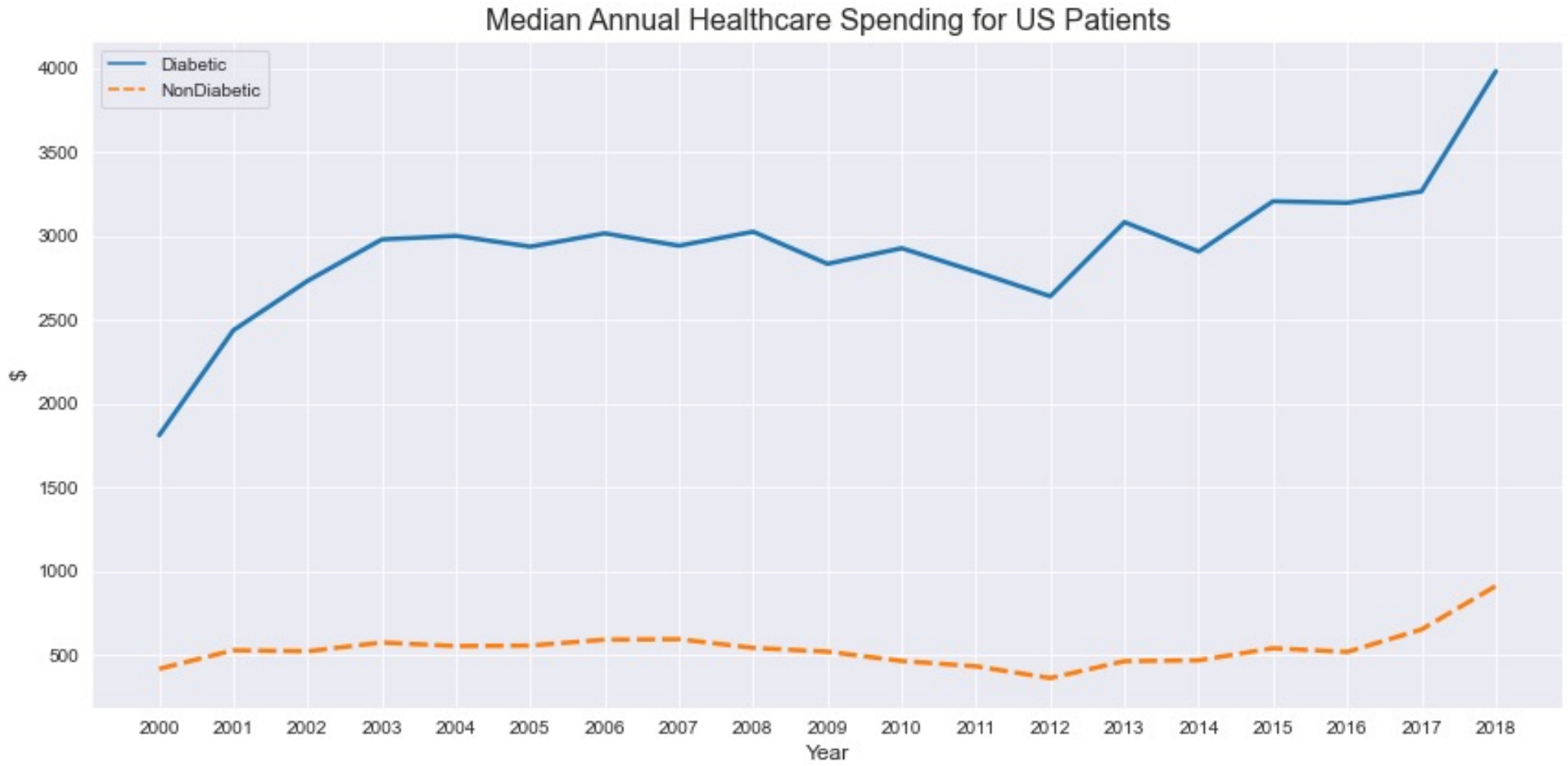
# DATA

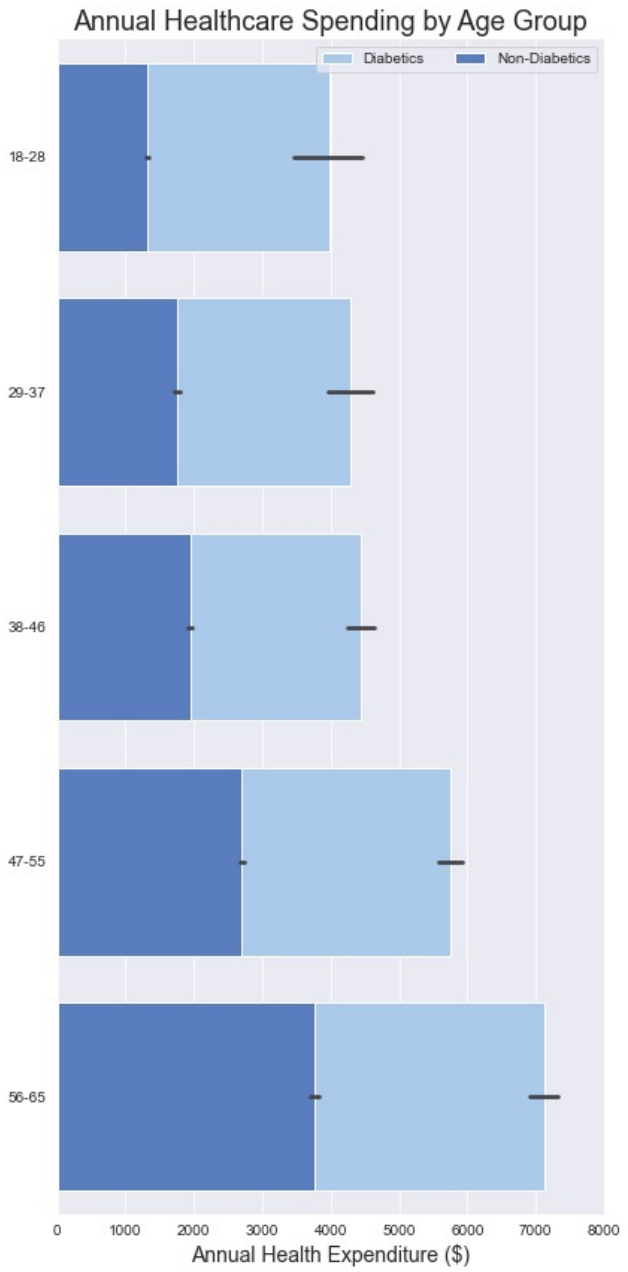
- Data in this project is sourced from the Medical Expenditure Panel Survey (MEPS), a nationally representative survey of the civilian noninstitutionalized population administered by the Agency for Healthcare Research and Quality.
  - Specifically, this project's data is pulled from MEPS's annual Household Full Year Consolidated Data Files.

# DATA

- Input data for modeling includes a survey subject's:
  - US geographic region
  - Age
  - Sex
  - Race/ethnicity
  - Education level
  - Degree of annual income
- The target data is the subject's annual healthcare expenses.

# VISUALIZATIONS & INSIGHTS






# VISUALIZATIONS & INSIGHTS

- For both groups, healthcare spending seems to increase as a person ages
- For diabetics, though, the increase in spending becomes more significant – especially for ages 38-46 and beyond
  - Compare differences between light blue boxes to differences between dark blue boxes

# VISUALIZATIONS & INSIGHTS

- Further insights were provided by classification, which was performed on 3 different groups:
  - **General**: Included data on both diabetics & non-diabetics. This model's purpose was to let me know how important having diabetes is in determining which healthcare expenditure tier a subject is part of.
  - **Diabetics**: Only included data on diabetic survey subjects. This model's purpose was to let me know which demographic features are most important in determining which healthcare expenditure tier a diabetic subject is part of.
  - **Non-Diabetics**: Only included data on non-diabetic survey subjects. This model's purpose was to let me know which demographic features are most important in determining which healthcare expenditure tier a non-diabetic subject is part of.

# VISUALIZATIONS & INSIGHTS

- Ultimately, these were the results for each:
  - **General**: whether a subject had diabetes was the 3rd most important variable in determining which healthcare expenditure tier they were in.
  - **Diabetics**: demographic variables that were significant in determining healthcare expenditure tier– subject's age
  - **Non-Diabetics**: demographic variables that were significant in determining healthcare expenditure tier – subject's age, sex, income

# CONCLUSIONS

- To answer the business problem's central question, the most impactful factor on healthcare spending for both groups is a person's age.
- As seen in visualizations, though, healthcare spending increases more rapidly over time for diabetics.
  - This is likely due to the increased likelihood of medical issues as a person ages.
  - What might be a normal aging-related issue could be exacerbated for someone with an existing condition, like diabetes.

# CONCLUSIONS

- Based on this conclusion, I recommend that Beyond Type 1:
  - Direct fundraising efforts toward poorer middle-age diabetics, whose income might not keep up with rising healthcare costs over time.
  - Make an organization-wide push to recruit more middle-age diabetics into their advocacy efforts.
  - Create education programs for older citizens that ensure they maintain healthy habits & keep them up to date on the latest diabetes management technology.

# FUTURE WORK

- Given more time with the job, I would:
  - Obtain similar data that has a more thorough time component in order to perform time series based speculation about future healthcare costs.
  - Obtain healthcare spending data for US residents with other specific conditions to observe how much they spend compared to diabetics.
  - Make comparison of diabetes healthcare costs for US residents vs residents of other nations around the world.

# QUESTIONS?

Github: https://github.com/dl-gd
Email: dlgarcia.017@gmail.com