Team 2

Importance of Healthy Habits

Background

Health is crucial to everyone

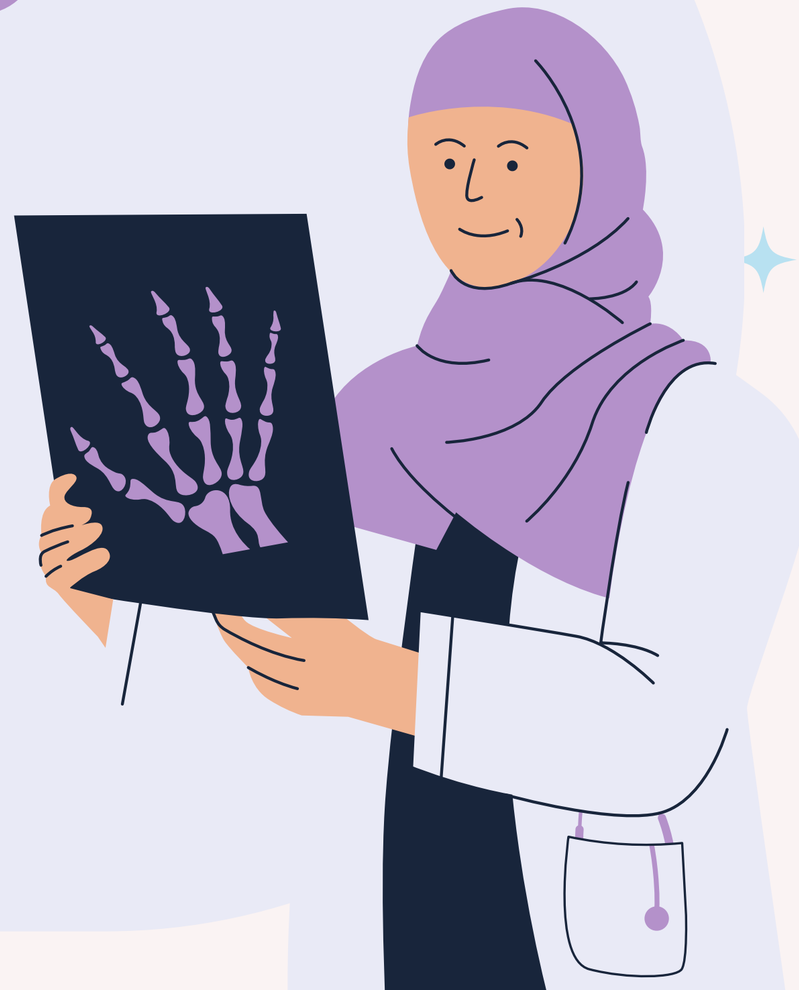
- FACTS: **Heart disease, cancer, and accidents** are top 3 causes of death in the US. In 2022, heart disease alone was responsible for over 700,000 deaths*.

- GOAL: Discovering the causality between personal habits & physical condition and chronic diseases.

*2022 statistics from Centers for Disease Control and Prevention(CDC)

Main Topics

Chronic Diseases	~	Habits / Physical Condition
1 Diabetes		Exercise Obesity
2 Hypertension		Obesity Smoker Sleep Time



Data Source

National Health and Nutrition Examination Survey (NHANES)

- Conducted yearly by the Centers for Disease Control and Prevention (CDC)
- Assess the health and nutritional status of adults and children in the United States

- Selected Period: **2017-2018**
- **9,254** participants completed the interview
- **8,704** participants were examined

Variables

Demographic

household Income
Gender
Age

Dietary

Total nutrients intake, consumed by participants during the 24-hour period

Saturated fatty acids (gm) ≤20
Sugars (gm) ≤50
Sodium (mg)≤2000
Fiber (gm) ≥31

Examination

BMI, blood pressure, pulse, Waist circumference, Hip circumference

Laboratory

Direct HDL-Cholesterol (mg/dL), under 60 mg/dL indicates the risk for hypertension

Cotinine, Serum (ng/mL), over 11 ng/mL can be recognized as smoker

Questionnaire

Sleep hours, exercise habits, diabetes
(mental health, self-report dietary habits, smoking habits, hypertension)

Methodology

Potential Outcome Frameworks

Observational study focusing on **ATE** - to understand the treatment effect in real world

Ensure unconfoundedness and overlap assumption

1. Choosing confounders based on credible sources
2. Confirm that all models meet overlap assumption requirements to conclude that the treatment effect is unbiased and valid.

Model Comparisons

Utilizing different models to ensure the validity of the suggestion.

Applied Models

Regression Based (w/o and w/ interaction)

Propensity-Score Based

Doubly Robust (Augmented IPW)

Doubly Robust PLM

Causal Analytics

Hypothesis: Diabetes & Exercise

	Regression Based	Regression Based (w/ Interaction)	IPWE	AIPWE	PLM
ATE Estimate	-0.2549	-0.2319	-0.0356	-0.0309	-0.0301
Standard Error	0.115	0.146	0.014	0.016	0.009
t-statistic	-2.22	-1.59	-2.54	-1.93	-3.34

***Target: individuals aged 21 or older**

****Control covariates: Alcohol, Smoke, Sleep, BMI, Dietary, Income, Age, Gender, etc**

*****Number of observed units: 3,275**

Causal Analytics

Hypothesis: Diabetes & BMI

	Regression Based	Regression Based (w/ Interaction)	IPWE	AIPWE	PLM
ATE Estimate	0.7941	0.8016	0.1033	0.0945	0.1032
Standard Error	0.116	0.139	0.0148	0.0130	0.0094
t-statistic	6.865	5.766	6.980	7.269	10.979

***Target: individuals aged 21 or older**

****Control covariates: Exercise, Alcohol, Smoke, Sleep, Dietary, Income, Age, Gender,Pulse, Blood pressure**

*****Number of observed units: 2,534**

Causal Analytics

Hypothesis: Hypertension & BMI

	Regression Based	Regression Based (w/ Interaction)	IPWE	AIPWE	PLM
ATE Estimate	-4.8640	-5.1344	-6.1644	-5.0379	-4.9365
Standard Error	0.597	0.637	0.5632	0.4320	0.3298
t-statistic	-8.149	-8.06	-10.9460	-11.6622	-14.9679

***Hypertension is measured by the HDL-C(High-density lipoprotein cholesterol) level, normally this value is larger than 60mg/dL.**

****Control covariates: Exercise, Blood pressure, Alcohol, Smoke, Sleep, Dietary, Income, Age, Gender, etc**

*****Number of observed units: 3,244**

Causal Analytics

Hypothesis: Hypertension & Smoking

	Regression Based	Regression Based (w/ Interaction)	IPWE	AIPWE	PLM
ATE Estimate	-1.3546	-1.4727	-2.0406	-2.0403	-1.4970
Standard Error	0.4957	0.5315	0.6530	0.5667	0.5130
t-statistic	-2.7327	-2.7706	-3.1249	-3.6001	-2.9160

*Smoking is measured by the indicator of Cotinine >= 11 ng/mL
**Control covariates: Alcohol, BMI, Dietary, Sleep, Exercise, Income, Gender, Age
***Number of observed units: 3,817

Causal Analytics

Hypothesis: Hypertension & Sleep

	Regression Based	Regression Based (w/ Interaction)	IPWE	AIPWE	PLM
ATE Estimate	-0.9243	-0.8480	-0.5613	-0.5759	-0.3647
Standard Error	0.578	0.608	0.6066	0.5385	0.3389
t-statistic	-1.600	-1.395	-0.9253	-1.0694	-1.0761

***Sleep well is measured by the indicator of sleep hour >= 7 hour on weekdays**

****Control covariates: Alcohol, Exercise, Smoke, BMI, Dietary, Income, Age, Gender, etc**

*****Number of observed units: 3071**

Conclusion

You ate your way in! You can walk your way out!

- Obesity is significant to diabetes and hypertension.
- Smoking is a significant contributor to hypertension
- Although there are studies indicating lack of sleep is relevant to hypertension, this dataset has not shown significant causality.

Limitation & Future Research

- Limits of observational studies
- Expand study timeframe to more recent data
- Researching on more potential factors

Thank you for your attention