

Using Data to Identify Obesity Rates in the U.S.A

Group Final Project ADTA 4240.700

By:

Hildalis Diaz, Robert Knutzen, Kyle Kildare, Linda
Bayet, Mohd Maher, Rachel Lara

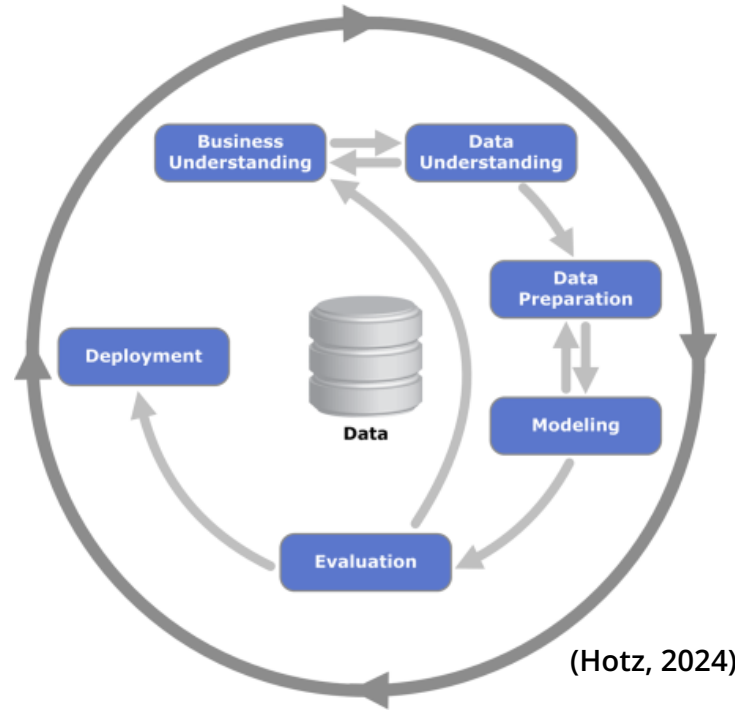
Problem Statement

It is no question that obesity is a huge problem within the United States. Adult obesity rates have been increasing since 1990. In fact, according to the Global Obesity Observatory, the United States currently ranks 19th (adults) and 22nd (children) out of 200 countries. This includes both male and female genders as well as every age group and ethnicity in every state in the country.

link with more information: <https://data.worldobesity.org/rankings/?age=c&sex=t>



Data Analytics Life Cycle Using CRISP-DM Framework



(Hotz, 2024)

Business Understanding

The Administration for Children and Families (ACF) works with states, tribes, and community organizations to promote the well-being of families, children, and communities. They would like to hire us to evaluate obesity rates around the country. Essentially, they would like to look into communities with the highest obesity rates depending on demographic groups. Their goal is to lower obesity rates, and will use the data provided by us to find where an outreach fitness program is most needed.

The administration has a budget specific for the program. Our data would assist in deciding how their resources should be allocated according to need, with the hope to institutionalize a habit of fitness and ultimately tackle obesity.

Data Modeling

- Retrieve CSV File Via Web
- Clean In OpenRefine
- Create Table/Run In BigQuery
- Use Looker To Visualize Data
- Evaluated Visualized Data For Business Needs



Data Preparation with OpenRefine


Using
OpenRefine
we're able to
clean the raw
data using
features
available in
OpenRefine.

OpenRefine Nutrition Physical Activity and Obesity Behavioral Risk Factor Surveillance System csv [Permalink](#) [Open...](#) [Export](#) [Help](#)

Facet / Filter Undo / Redo 0 / 6 **104,272 rows** Extensions Wikibase

Show as: rows records Show: 5 10 25 **50** 100 500 1000 rows

« first « previous 1 - 60 next » last »

Using facets and filters 
Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.
Not sure how to get started? [Watch these screencasts](#)

	All	Year	Year	Loca	Loca	Data	Class	Topic	Question	Data	Data	Data	Data	Data	Data
1.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have obesity	2011	Value	23.3	23.3		
2.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have obesity	2011	Value	35.5	35.5		
3.	2011	2011	AK	Alaska	BRFSS	Physical Activity	Physical Activity - Behavior	Physical Activity - Behavior	Percent of adults who achieve at least 150 minutes a week of moderate-intensity aerobic physical activity or 75 minutes a week of vigorous-intensity aerobic physical activity and engage in muscle-strengthening activities on 2 or more days a week	2011	Value	21.7	21.7		
4.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have obesity	2011	Value	29.5	29.5		
5.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have an overweight classification	2011	Value	33.5	33.5		
6.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have obesity	2011	Value	28.0	28.0		
7.	2011	2011	AK	Alaska	BRFSS	Physical Activity	Physical Activity - Behavior	Physical Activity - Behavior	Percent of adults who engage in muscle-strengthening activities on 2 or more days a week	2011	Value	28.8	28.8		
8.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have an overweight classification	2011	Value	38.7	38.7		
9.	2011	2011	AK	Alaska	BRFSS	Physical Activity	Physical Activity - Behavior	Physical Activity - Behavior	Percent of adults who engage in muscle-strengthening activities on 2 or more days a week	2011	Value	35.9	35.9		
10.	2011	2011	AK	Alaska	BRFSS	Physical Activity	Physical Activity - Behavior	Physical Activity - Behavior	Percent of adults who achieve at least 300 minutes a week of moderate-intensity aerobic physical activity or 150 minutes a week of vigorous-intensity aerobic activity (or an equivalent combination)	2011	Value			~	Data not
11.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have an overweight classification	2011	Value			~	Data not
12.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have obesity	2011	Value	26.7	26.7		
13.	2011	2011	AK	Alaska	BRFSS	Physical Activity	Physical Activity - Behavior	Physical Activity - Behavior	Percent of adults who achieve at least 150 minutes a week of moderate-intensity aerobic physical activity or 75 minutes a week of vigorous-intensity aerobic physical activity and engage in muscle-strengthening activities on 2 or more days a week	2011	Value			~	Data not
14.	2011	2011	AK	Alaska	BRFSS	Physical Activity	Physical Activity - Behavior	Physical Activity - Behavior	Percent of adults who achieve at least 150 minutes a week of moderate-intensity aerobic physical activity or 75 minutes a week of vigorous-intensity aerobic physical activity and engage in muscle-strengthening activities on 2 or more days a week	2011	Value	17.3	17.3		
15.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have obesity	2011	Value	25.8	25.8		
16.	2011	2011	AK	Alaska	BRFSS	Obesity / Weight Status	Obesity / Weight Status	Obesity / Weight Status	Percent of adults aged 18 years and older who have an	2011	Value	32.0	32.0		

Data Preparation with OpenRefine: Removing Columns

ADTA 4240

GeoLocation	Class
(64.845079957001, -147.722059036)	OWS
(64.845079957001, -147.722059036)	OWS
(64.845079957001, -147.722059036)	PA

93,505 matching rows (104,272 total)

Show as: rows records Show: 5 10 25 50 100 5

All	Year	Loca	Loca	Data	Cl
Transform...	2011	AK	Alaska	BRFSS	Obesit
Edit all columns	2011	AK	Alaska	BRFSS	Status
Facet	2011	AK	Alaska	BRFSS	Physic
Add blank rows					
Edit rows					
Edit columns					
View					
6.	2011				
7.	2011	2011	AK	Alaska	BRFSS
8.	2011	2011	AK	Alaska	BRFSS

Re-order / Remove columns

Drag columns to re-order

Drop columns here to remove

Obesity_Usage	
Total	
Age(years)	
Education	
Sex	
Income	
Race/Ethnicity	
ClassID	
TopicID	
QuestionID	
DataValueTypeID	
LocationID	
StratificationCategory1	
Stratification1	
StratificationCategoryID1	
StratificationID1	

Remove all Add all

OK Cancel

Data Preparation with OpenRefine: Removing Data and Exporting Clean Data

ADTA 4240

OpenRefine Nutrition Physical Activity and Obesity Behavioral Risk Factor Surveillance System csv [Permalink](#) [Open...](#) [Export](#) [Help](#)

Facet / Filter Undo / Redo 5 / 5 104,272 rows Extensions Wikibase

Show as: rows records Show: 5 10 25 50 100 500 1000 rows < first < previous 1 - 60 next > last >

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started?
[Watch these screencasts](#)

tion	Data	Data	Data	Data	Data	Data_Value_Footnote	Low	High	Sam	Total	Age	Education	Sex
Adults aged 18 years and older who have obesity	2011	Value	28.0	28.0			24.9	31.3	1620				Male
Adults who engage in muscle-strengthening on 2 or more days a week	2011	Value	28.8	28.8			21.9	36.8	232			Less than high school	
Adults aged 18 years and older who have an "A" classification	2011	Value	38.7	38.7			32.7	45.2	451		25 - 34		
Adults who engage in muscle-strengthening on 2 or more days a week	2011	Value	35.9	35.9			33.2	38.6	2417				
Adults who achieve at least 300 minutes a week of intensity aerobic physical activity or 150 minutes a vigorous-intensity aerobic activity (or an equivalent on)	2011	Value			~	Data not available because sample size is insufficient.							
Adults aged 18 years and older who have an "A" classification	2011	Value			~	Data not available because sample size is insufficient.							
Adults aged 18 years and older who have obesity	2011	Value	26.7	26.7			23.9	29.7	1716				Female
Adults who achieve at least 150 minutes a week of intensity aerobic physical activity or 75 minutes a vigorous-intensity aerobic physical activity and muscle-strengthening activities on 2 or more days	2011	Value			~	Data not available because sample size is insufficient.							
Adults who achieve at least 150 minutes a week of intensity aerobic physical activity or 75 minutes a vigorous-intensity aerobic physical activity and muscle-strengthening activities on 2 or more days	2011	Value	17.3	17.3			12.5	23.4	349				!
Adults aged 18 years and older who have obesity	2011	Value	25.8	25.8			22.3	29.6	954			Some college or technical school	
Adults aged 18 years and older who have an "A" classification	2011	Value	32.0	32.0			24.7	40.4	212		18 - 24		
Adults who engage in no leisure-time physical	2011	Value	19.2	19.2			15.1	24.1	530				!
Adults who achieve at least 150 minutes a week of intensity aerobic physical activity or 75 minutes a vigorous-intensity aerobic activity (or an equivalent on)	2011	Value	56.2	56.2			44.9	66.9	130				
Adults who engage in no leisure-time physical	2011	Value			~	Data not available because sample size is insufficient.							
Adults aged 18 years and older who have an "A" classification	2011	Value	31.3	31.3			28.1	34.6	1716				Female

BigQuery Data

console.cloud.google.com/bigquery?hl=en&project=e-analogy-448500-r4&ws=11m10!1m4!1m3!1se-analogy-448500-r4!2sbqxjob_1309e0...

Google Cloud myfirstproject2025MMaher Search (/) for resources, docs, products, and more Search

BigQuery

Studio

Pipelines & Integration

- Data transfers
- Pipelines (Dataform)
- Scheduled queries
- Scheduling **Preview**

Governance

- Analytics Hub
- Policy tags

Administration

- Monitoring
- Jobs explorer
- Partner Center
- Settings **Preview**
- Release Notes

Untitled query

```

1 select *
2 from 'e-analogy-448500-r4.ADTA.Obesity';
3
4
5 select YearStart, LocationAbbr, Class,
6 Data_value, Data_value_Alt, Sex, YearEnd, Stratification1
7 from 'e-analogy-448500-r4.ADTA.Obesity'
8 where class = "Obesity and Weight Status" and Data_Value > 50
9 order by YearStart;
  
```

Query completed

Query results

SAVE RESULTS OPEN IN

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS		EXECUTION GRAPH	
Row	YearStart	LocationAbbr	Class	Data_value	Data_value_Alt	Sex	YearEnd	Stratifica
1	2011	NV	Obesity and Weight Status	59.2	59.2	null	2011	America
2	2011	LA	Obesity and Weight Status	50.5	50.5	null	2011	2 or mon
3	2011	NM	Obesity and Weight Status	53.4	53.4	null	2011	2 or mon
4	2011	VT	Obesity and Weight Status	50.4	50.4	null	2011	Hispanic
5	2011	NV	Obesity and Weight Status	54.6	54.6	null	2011	Other
6	2012	HI	Obesity and Weight Status	51.4	51.4	null	2012	Hawaiian
7	2012	AL	Obesity and Weight Status	51.4	51.4	null	2012	2 or mon
8	2012	DE	Obesity and Weight Status	50.2	50.2	null	2012	Asian

Results per page: 50 1 - 50 of 106

Data Understanding

ADTA 4240

Looker Studio Reporting - 2/27/25, 5:32 PM

File Edit View Insert Page Arrange Resources Help

+ Add quick filter

Reset

projectdata1

LocationAbbr	YearStart
1. VI	1820
2. HI	1613
3. NV	1209
4. WI	1009
5. AR	1010
6. ND	8081
7. MO	8070
8. UT	6060
9. SD	6057
10. WV	6055
11. ME	6054

Let's get started

Drag a field from the Data Panel to the canvas to add a new chart or select a component on the report canvas to edit it.

Data

Search

projectdata1 - 2/27/25, 5:32 PM

Age, years

Class

ClassID

Data, Value

Data, Value, Alt

Data, Value, Endnote

Data, Value, Endnote, Symbol

Data, Value, Type

Data, Value, Unit

Demographic

Demographic, SpecID

Education

Sex, Location

High, Confirmed, Limit

Income

LocationAbbr

LocationClass

LocationID

Low, Confirmed, Limit

Question

Add a field

Add a parameter

Add a data

Class Obesity and Weight Status (1) X

Reset

Record Count

106

Data, value

5,671

Class

1. Obesity and Weight Status

1 - 1/1

Record Count

20

10

0

2011

2012

2013

2014

2015

2016

2017

2018

2019

2020

2021

2022

2023

Stratification1

1. American Indian/Alaska Native

2. 2 or more races

3. Non-Hispanic Black

4. Other

5. Hawaiian/Pacific Islander

6. 45 - 54

7. Hispanic

6.6%

9.4%

10.4%

11.3%

15.1%

35.8%

American Indian/Alaska Native

2 or more races

Non-Hispanic Black

Other

Hawaiian/Pacific Islander

45 - 54

Hispanic

Less than \$15,000

\$75,000 or greater

Let's get started

Drag a field from the Data Panel to the canvas to add a new chart or select a component on the report canvas to edit it.

Data

Search

BigQuery Custom SQL...

Class

Data, value

Data, value, Alt

LocationAbbr

Sex

Stratification1

YearEnd

YearStart

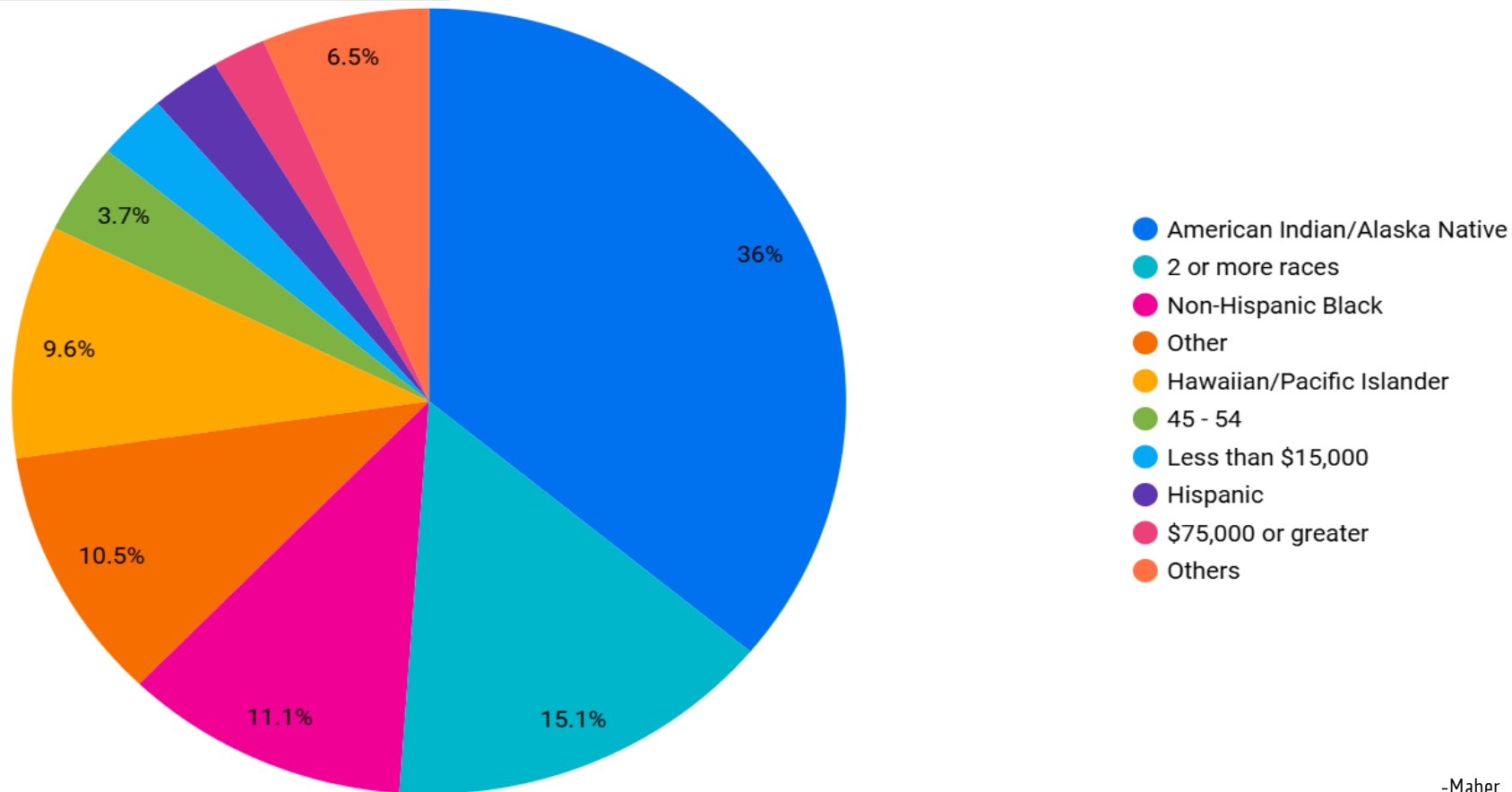
Record Count

Add a field

Add a parameter

Add a data

Different Demographic by Obesity Rate



Data Evaluation

- Start with a clear understanding of what the client is trying to achieve
- Appropriate selection of data tools
- Correct method of evaluation for data type (quantitative data)
- Cleaning and organizing the data is vital
- Analyze the data and interpret it for the customer



References:

ADTA 4240

-Diaz,Knutzen

Center for Obesity Research and Education. (n.d.). College of Public Health.
<https://cph.temple.edu/research/research-centers/center-obesity-research-education>

Childhood obesity is a complex health issue. (2022, July 15). Centers for Disease Control and Prevention.
<https://www.cdc.gov/obesity/basics/consequences.html>

Hotz, N. (2024, December 9). What is CRISP DM? Data Science PM.
<https://www.datascience-pm.com/crisp-dm-2/>

Ranking (% obesity by country). (n.d.). World Obesity Federation Global Obesity Observatory. <https://data.worldobesity.org/rankings/>

Assistant Secretary for Public Affairs (ASPA). (2024). *HHS Agencies & Offices*. HHS.gov.
<https://www.hhs.gov/about/agencies/hhs-agencies-and-offices/index.html>



Team Members & Work breakdown by slide

Linda Bayet: Slide 3

Hildalis Diaz: Slides 1, 2, 4, 13

Kyle Kildare: Slides 4, 5, 10, 12. Editing, formatting, and graphics, all slides.

Robert Knutzen: Slides 4, 5, 10, 12, 13. Editing and graphics on all slides.

Rachel Lara: Slides 4, 6, 7, 8.

Mohd Maher: Slides 11 and 12.