Consumer Demographic Data Analysis

Randall C. Crawford

MS Data Science, National University

ANA 600: Fundamental of Analytics

Matthew C. Vanderbilt, MSBA

October 6, 2024

This was to be written as a full APA report as if you were submitting to your supervisor. You've contained lists of information but have limited interpretations and incomplete paragraphs.

Consumer Demographic Data Analysis

This data was collected to gather information about consumer attitudes and decisions to save, borrow, or make purchases. The researchers wanted to forecast changes in consumer behavior in the United States and better understand consumer confidence about personal finances, employment, price changes, and the perceived state of national business. Below is a list of the variables included in the dataset of 2,000 observations.

Variable	Description/Question
household	Including yourself, how many members of your household are 18 years or older?
kids	How many members of your household are 17 years or younger?
vehicles	How many vehicles do you use, including those leased, owned, or provided by your employer for personal use?
priceExpected	By what percent do you expect prices to go up/down on average during the next 12 months?
incomeExpected	By what percent per year do you expect your (family) income to (increase/decrease) during the next 12 months?
businessExpected	Considering business conditions in the country as a whole, do you think that during the next 12 months, we'll have good times or bad times financially?
financialStability	Would you say that you (and your family) are better off or worse off financially than you were a year ago?
investments	What type of investments do you plan to add or shift money into or open during the next three months: mutual funds, savings accounts, stocks, bonds, retirement accounts?
income	What is your household gross income? (in thousands)
age	What is your age?

employmentSector	What is your employment sector?
region	What is your state of residence? (recoded into regions)
hoursPerWeek	Approximately how many hours a week are you employed?

1. The quantitative variable **household** was separated into four family categories (***Traditional****: 2 adults,* ***Modern****: 3 or 4 adults,* ***Extended****: 5 or 6 adults,* ***Large****: 7 through 10 adults*). Observations were allocated across the four family categories (***Traditional****: 299,* ***Modern****: 1117,* ***Extended****: 508,* ***Large****: 76*).

Mean Values:

	hoursPerWeek	income	incomeExpected	priceExpected	vehicles	kids	age*
Traditional	38.92308	48.24415	1.217391	1.304348	1.331104	1.000000	35.47687
Modern	38.70367	47.02328	1.227395	1.186213	1.743062	1.916741	36.13521
Extended	38.23622	45.57677	1.236220	1.234252	2.025591	3.277559	35.81410
Large	38.43421	46.02632	1.078947	1.368421	2.407895	4.763158	35.49254

see https://apastyle.apa.org/style-grammar-guidelines/tables-figures/tables for proper formatting

Standard Deviation Values:

	hoursPerWeek	income	incomeExpected	priceExpected	vehicles	kids	age*
Traditional	6.504317	13.75544	1.0505185	2.582403	0.4713992	0.0000000	9.110061
Modern	6.056108	13.46825	0.9974168	2.522882	0.5348590	0.6594073	9.186068
Extended	5.995339	13.32376	1.0506442	2.464391	0.6259782	0.8139128	9.033465
Large	6.348408	11.73709	1.0925586	2.354242	0.7514897	1.0566415	9.432974

*The quantitative variable **age** had a **156** reduction in observations (**1844**), due to no age being provided.

2000 observations were spread evenly over four regions (**region**).

Midwest	Northeast	South	West
500	500	500	500

2000 observations over 11 defined employment sectors (**employmentSector**):

Construction	Education	Financial Services
92	86	150

Health services	Leisure and Hospitality	Manufacturing
202	251	137

Other	Professional and Business Services	Public Administration
235	158	241

Transportation and utilities	Wholesale and Retail Trade

239 209

Mean Values:

	hoursPerWeek	income	age*
Construction	38.40217	48.25000	36.89412
Education	37.40698	48.20930	38.42169
Financial Services	38.57333	42.42000	35.32593
Health Services	38.19802	45.91089	35.96667
Leisure and Hospitality	38.61355	45.19920	35.97059
Manufacturing	38.60584	49.56204	36.00781
Other	38.62979	48.12766	34.29680
Professional and Business Services	38.84810	46.13291	35.85616
Public Administration	38.97510	44.90041	36.02691
Transportation and Utilities	38.53975	51.50628	36.36792
Wholesale and Retail Trade	39.05263	45.52153	36.03590

Standard Deviation Values:

	hoursPerWeek	income	age*
Construction	6.023872	13.94229	9.953491
Education	6.239114	15.15886	11.031132
Financial services	6.056054	11.93454	8.522866
Health Services	5.787308	12.95311	9.328816
Leisure and Hospitality	5.811545	13.33567	9.448237
Manufacturing	7.195135	13.28968	9.500515
Other	5.643539	13.25202	8.489526
Professional and Business Services	6.005493	13.14161	7.993092
Public Administration	6.559297	13.89958	9.017960
Transportation and Utilities	6.000481	13.68643	9.456636
Wholesale and Retail Trade	6.335348	11.75801	8.674155

Inferences Drawn:

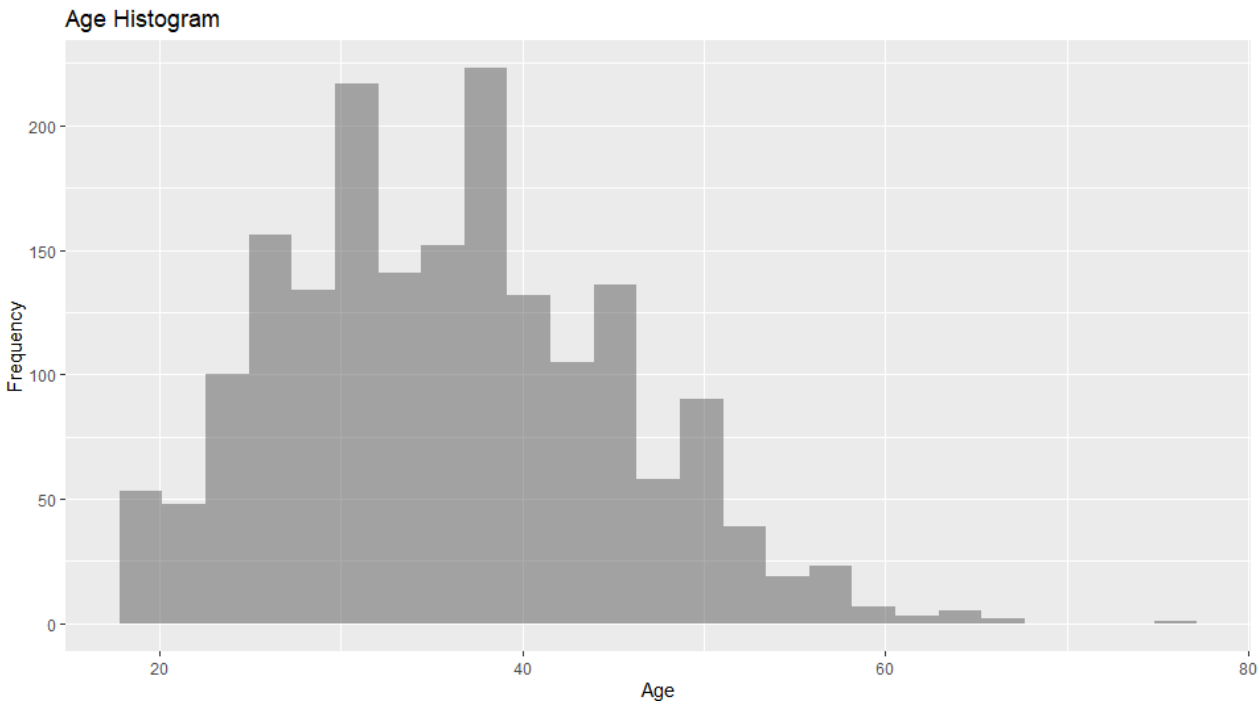
- The observations from **Traditional** (2 adult) family households had higher mean annual **income** and **hours per week**, probably due to the limited social dynamic allowing for more time and focus on a career.
- The observations for **Extended** (5 or 6) and **Large** (7 to 10) family households had a 3 to 5 **kid** mean, which suggest possibly more than one family living in the household.
- The observations for **Wholesale and Retail Trade** had a higher mean **hours per week** than other employment sectors, probably due to having a lower mean annual **income**.
- The observations for **Transportation and Utilities** had a higher mean annual **income** than other employment sectors.

- The observations for **Education** had a higher mean **age** than other employment sectors, and the non-standard **other** employment sector had the lowest mean **age**.
- It is concerning that **construction** and **manufacturing** had **229** total observations, while **1,536** observations reside in the service economy. A 30 year trend that needs to change.

I believe

2. Believe the intended population of observations in this study was reasonably good, touching on many demographic considerations. All participants were employed, and were part of a household with at least two adults. Mean and Standard Deviation calculations only reveal so much. I think Think further correlation analysis would be useful.

3. Given that there are many variables to consider, one must be careful with sample selection from this dataset to try to maintain meaningful representation, independence, and randomness. The **age** quantitative variable results were carefully separated to ensure no bias was introduced to other quantitative variables.



capitalize the beginning of a sentence

4. in the event of further data collection, it would be good to know how many employed and dependent adults reside in the household. It is hard to know if the case subject could be supporting a number of others or benefiting from living with others with their own income.