Calculations and results could be accessed in this spreadsheet.

Question 1:

Categories	Column Labels
Demographic	Gender, Marital Status, Work Status, Education, Annual Income, Age, Location
Purchasing	Purchasing Decision-maker, Purchasing Location, Purchasing frequency, Monthly Electronics Spend, Monthly Household Spend
Behavioral	TV Viewing[1]
Attitude	Technology Adoption
Other	Favorite feature

Question 2:

$$Annual\ Spending\ on\ Electronics\ =\ Monthly\ Electronics\ Spend\ \times 12$$

$$Spending\ as\ Percentage\ of\ Income\ =\ \frac{Annual\ Spending\ on\ Electronics\ (\$)}{Annual\ Income\ (\times 1000\$)}$$

Question 3:

Numerical variables are separated into intervals according to their distribution:

- Income distribution shows a clear clustering behavior, thus separated into two intervals: Below $40,000 and above $40,000.
- Age follows a uniform distribution ranging from 18 to 80, thus separated into five groups: 18-34, 35-44, 45-54, 55-64, and above 65.
- Monthly electronics spending is separated into three groups: below $20/month, $20 - $50/month and above $50/month
- Monthly household spending is separated into three groups: below $100/month, $100-$200/month, and above $200/month
- TV Viewing behavior is separated into three groups: 1hr and below, 2hr-5hr, 6hr and above

[1] I feel there is just a blur line between behavioral and purchasing. The wikipedia lists purchasing as part of behavioral segmentation.

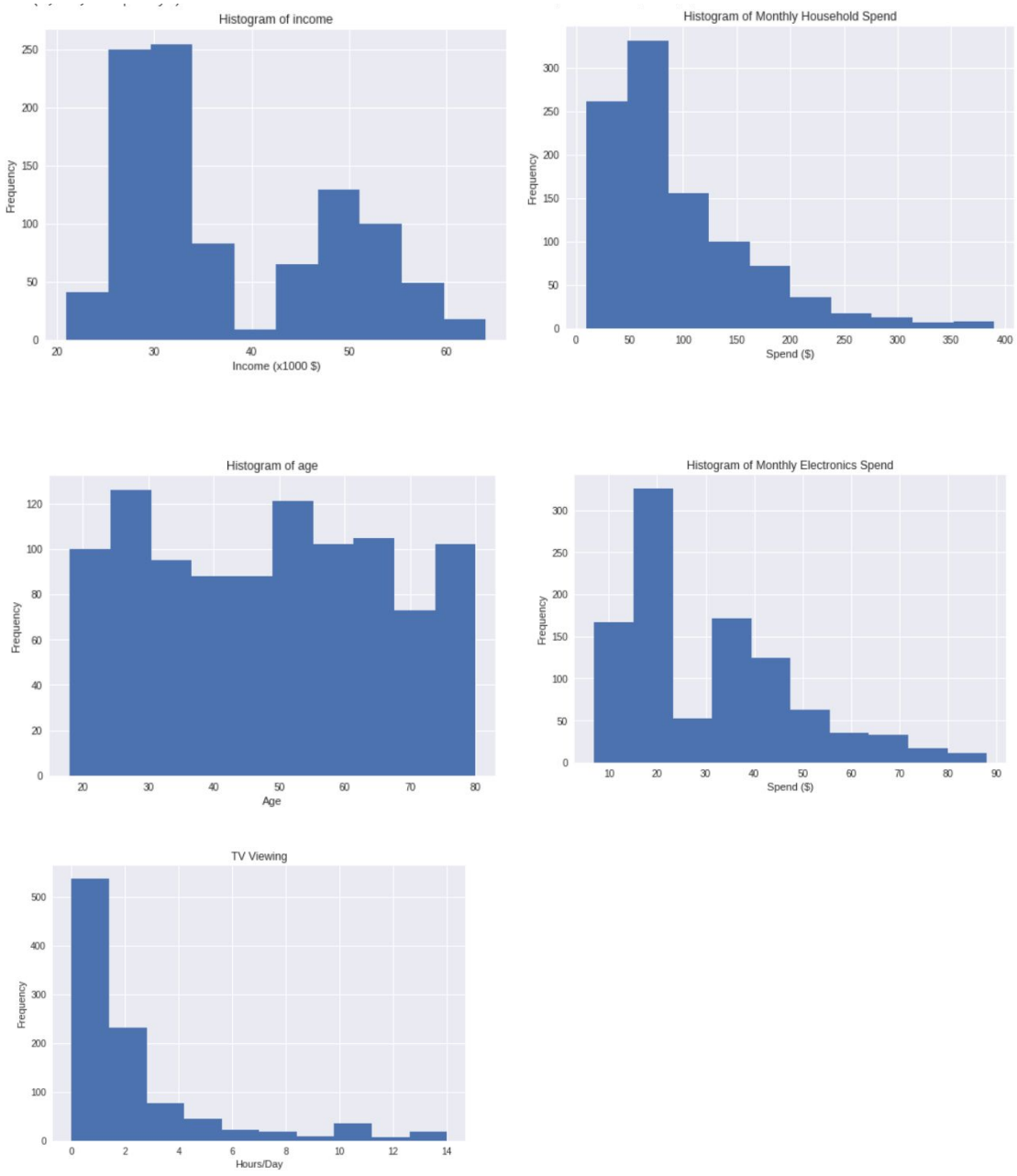

Fig 1: Histograms for numerical variables in survey data.[2]

[2] #dataviz: utilize graphs to find insights in the data

Gender	Male	535	53.5%
	Female	465	46.5%
Marital Status	Married	720	72.0%
	Single	280	28.0%
Work Status	None	350	35.0%
	Professional	650	65.0%
Education	None	506	50.6%
	MA	124	12.4%
	BA	260	26.0%
	PhD	110	11.0%
Purchasing decision maker	Single	440	44.0%
	Family	560	56.0%
Purchasing location	Specialty stores	170	17.0%
	Retail	294	29.4%
	Web	43	4.3%
	Discount	293	29.3%
	Mass-consumer electronics	200	20.0%
Technology adoption	Early	800	80.0%
	Late	200	20.0%
Favorite feature	Saving favorite shows to watch as a family	200	20.0%
	Time shifting	221	22.1%
	Schedule control	221	22.1%

	Cool gadget	228	22.8%
	Programming/interactive features	130	13.0%
Annual Income	< $40,000	630	63.0%
	>= $40,000	370	37.0%
Age	18-34	290	29.0%
	35-44	149	14.9%
	45-54	161	16.1%
	55-64	170	17.0%
	65+	230	23.0%
Monthly Electronics Spend	< $20	403	40.3%
	$20 - $50	452	45.2%
	>= $50	145	14.5%
Monthly Household Spend	< $100	669	66.9%
	$100 - $200	251	25.1%
	>= $200	80	8.0%
TV Viewing	<= 1hr	537	53.7%
	2hr - 5hr	353	35.3%
	>= 6hr	110	11.0%

Categorization by states are done externally on Python here[3].

Question 4:

a. How many married men who are early adopters can afford to purchase a TiVo for $499 and on average have enough money to purchase another electronic gadget in the next two years?

→ 134.

[3] Link also include code to preprocess and plot histograms from the data.

b. How many women with education of MA or PhD are making purchasing decisions for electronics without discussing them with a spouse, either because they are single, or because they are making purchasing decisions without the involvement of their spouses?

→ 55

c. Among early adopters, how many purchase electronics at least once every year and do so in stores that specialize in electronics?

→ 132

d. How many seniors (above the age of 65) spend more than six hours a day watching TV? What is their income range? What is their average annual income?

→ Seniors watching > 6 hrs of TV: 20 (65 yrs old and 6 hrs of watching movie is not counted)
→ Income range: $41000 to $55000
→ Average annual income: $48,600

Question 5:

→ Pearson's r: 0.1324652759
→ r^2: 0.01754704932

Question 6:

a. Code Gender as a number. What is the correlation between it and Annual income?

→ Pearson's r: 0.06747262269
→ r^2: 0.004552554813

b. Explain why it makes no difference which numbers are used to code Gender or other non-numeric attributes.

Because we standardize all values by converting them to deviation scores (i.e. subtract the mean of each variable from each value of that variable). As long as our coding rule is consistent (e.g. male always equal to 1, female always equal to 0), then the correlation r remain the same regardless of the value we encode the variable to.

Question 7:

a. Age and Purchasing	Pearson's r:	-0.001292617

frequency	r^2	1.67086E-06
b. Annual Income and TV Viewing	Pearson's r:	0.084335247
	r^2	0.007112434
c. Education and Favorite Feature	Pearson's r:	-0.048509391
	r^2	0.002353161
d. Monthly Electronics Spend and Monthly Household Spend	Pearson's r:	0.803805771
	r^2	0.646103717

Of the four pairs of attributes, only the monthly electronics spend - monthly household spend pair achieve high r^2 (0.65, much higher than the baseline 0.45 to be considered high). This means that the two variables are positively, linearly correlated to a high extent, which makes sense as electronics spending is part of the total household spending.

Also notice that Pearson's r does not indicate non-linear correlation. If two variables are not linearly correlated (e.g. their scatter plot could resemble a quadratic function), Pearson's r could still be low even though the two variables are actually related. Furthermore, for non-binary non-numeric variables, such as favorite feature, where categories are not orderable, the Pearson's r is not as meaningful. [4]

Question 8:
In this section, I abide by the constraint in the number of section impose by the question (7 segments in scheme 1 and 4 segments in scheme 2).

A. Scheme 1: 7 segments

	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Segment 7
Market size (% of TV-involved	14.8%	16.2%	2.1%	10.8%	24.8%	20.2%	11.1%

[4] #correlation: Calculate and critique correlation results.

households)							
Average annual income (x1000 $)	29.87162162	29.96914	51	50.09259	29.75806	57.57426	38.1982
Description of segment	Low Income Young Single	Low Income Young Married	High Income Young Single	High Income Young Married	Low Income Old Married	High Income Old Married	Old Single (Both Low and High Income)
Most appealing feature/benefit	Programming/Interactive features (81.1%)	Time shifting (35.8%)	Saving favorite shows to watch as a family (52,4%)	Saving favorite shows to watch as a family (47.2%)	Time Shifting (31.9%)	Saving favorite shows to watch as a family (63.9%)	Cool Gadget (29.7%)
Stores shopped for electronics	Retail (35.1%)	Retail (51.9%)	Mass-consumer electronics (52.4%)	Mass-consumer electronics (52.4%)	Retail (48.4%)	Mass-consumer electronics (63.9%)	Retail (34.2%)
Average electronics purchase per year ($)	360.24	218.96	593.71	600	202.54	568.21	364.64

B. Scheme 2: 4 segments

Segment Name	Segment 1	Segment 2	Segment 3	Segment 4
Market size (% of TV-involved households)	29.4%	29.3%	20.0%	21.3%
Average annual income (x1000 $)	$29,877.00	$29,979.52	$48,095.00	$55,492.96
Description of segment	People who usually shop for electronics from retail	People who usually shop for electronics from discount	People who usually shop for electronics from mass-consumer electronics	People who usually shop for electronics from specialty stores and web

Most appealing feature/benefit	Schedule control: 28.91%	Cool gadget: 29.35%	Saving favorite shows to watch as a family: 100%	Cool gadget: 27.23%
Stores shopped for electronics	Retail	Discount	Mass-consumer electronics	Specialty stores and web
Average electronics purchase per year ($)	$234.20	$228.86	$496.86	$639.72

Comments:

The first scheme (7 segments) shows problems of over segmentation. The segments are not externally heterogeneous. For example, segments 3, 4, and 6 share the same characteristics (including average income, average electronics spending, most appealing feature, and stores shopped for electronics). Segment 7 is not internally homogeneous as it includes both low income and high income participants, whose needs and behaviors are different from each other. This explains the low percentage of the most dominant feature and shopping channel, compared to other segments. It is not without merits, however. We recognize that only well-off customers (whose income is above $40,000 a year) could barely afford to buy the PVR up front. The market share of these users together is relative high (37.7%), and could be an important factor in deciding targets for marketing campaigns.

The second scheme (4 segments) is reasonably homogeneous internally, and heterogeneous externally. We might want to target segment 3, as they are wealthy enough to afford the PVR, while homogeneous in shopping pattern (i.e. 100% shop at mass-consumer electronics shops). These users also share a similar feature, namely saving favorite shows to watch as a family. It is likely that these users would relate to the same identities - connection, belonging, and heirloom. Thus, it would be easier to reach out to them than to other segments, and technological adoption attitude

Analysis:

Market research plays an utmost important role in product launching. Particularly, the process should reveal patterns in behaviors, attitudes and desires that the company could leverage to come closer to its potential users.

In this scenario, the research conducted by ThinkAlike gathers primary data directly concerning TiVo's product. The data are holistic and reflect various aspects of customer behaviors. The

number of participants is reasonable (1000); the interview quality is high (30 minutes of phone call). Demographic information shows that the data are collected from a socio-geo-economically diverse population (states, incomes, ages). All data are quantifiable, thus easy to do inference on.

The data collected reflect what TiVo's product is and revolve around values that TiVo targets. As they provide a novel device intended for the regular American who is unfamiliar with complicated solutions such as the contemporary VCR, as well as a versatile solution that lets users interact with television programs in ways that are not available before, TiVo would definitely be interested to learn about customers' attitude towards new technologies, as well as features most desired by customers. Information of customers' income is helpful for the company to gauge whether its pricing is reasonable for average users. The data set overall would be central in structuring marketing messages that most resonate with each respective segment (e.g. highlight its direct access feature to people who want to save programs to watch with family or highlight its affordability and versatility to people who keep tight electronics spending budget). Also, the company could consider altering the product's name depending on the market segment it reaches out to. For example, Digital Video Recorder (DVR) would sound more attractive to people favoring cool gadgets, while Personal Video Recorder (PVR) would sound more user-friendly, staying true to the intended image of the device.

Nevertheless, the research process is not without any drawbacks. The data collected are purely quantitative, following a chiefly analytic thinking process (i.e. questions are inspired by features of the PVR/DVR only). The researcher could have followed up with more open ended questions to gather personal insights.

For example, while the survey does ask about participants' favorite feature, the response is limited to one of the five choices only. The interpretation of some of the choices is also vague - What does it even mean to be a cool gadget? Is it the design or the functionality that determines the "coolness" of the product? The question also does not explore the reason why the feature is desirable to the potential users. Different favorite features could actually point to a similar desire. An office worker would like to record programs as their inflexible schedule does not allow them to watch it when they want. Another follow-up question might show that they do not want to be an office worker in the first place. A teenager would like to show off a cool gadget as a way to express their identity. They might later reveal their dream of being a tech innovator. Both point to the same desire of self-actualization, which TiVo could focus on as their main marketing theme. Segmentation by identity based on such insights gathered from open-ended questions as above is particularly helpful for the company. This establishes a unique brand for TiVo that is easily resonated to by potential customers and is more difficult to copy by competitors (which is the VCR in this case). There is also more benefits when it comes to marketing messages. Presenting potential customers with complex lists of its multiple features would not help build the product's easy-to-install, user-friendly image. Instead, by focusing on one or two identities

that matter to the segment, TiVo could attach specific emotions to its product, which makes it distinguishable.[5]

[5] #critique: Thoroughly critique the design and result of the market research.