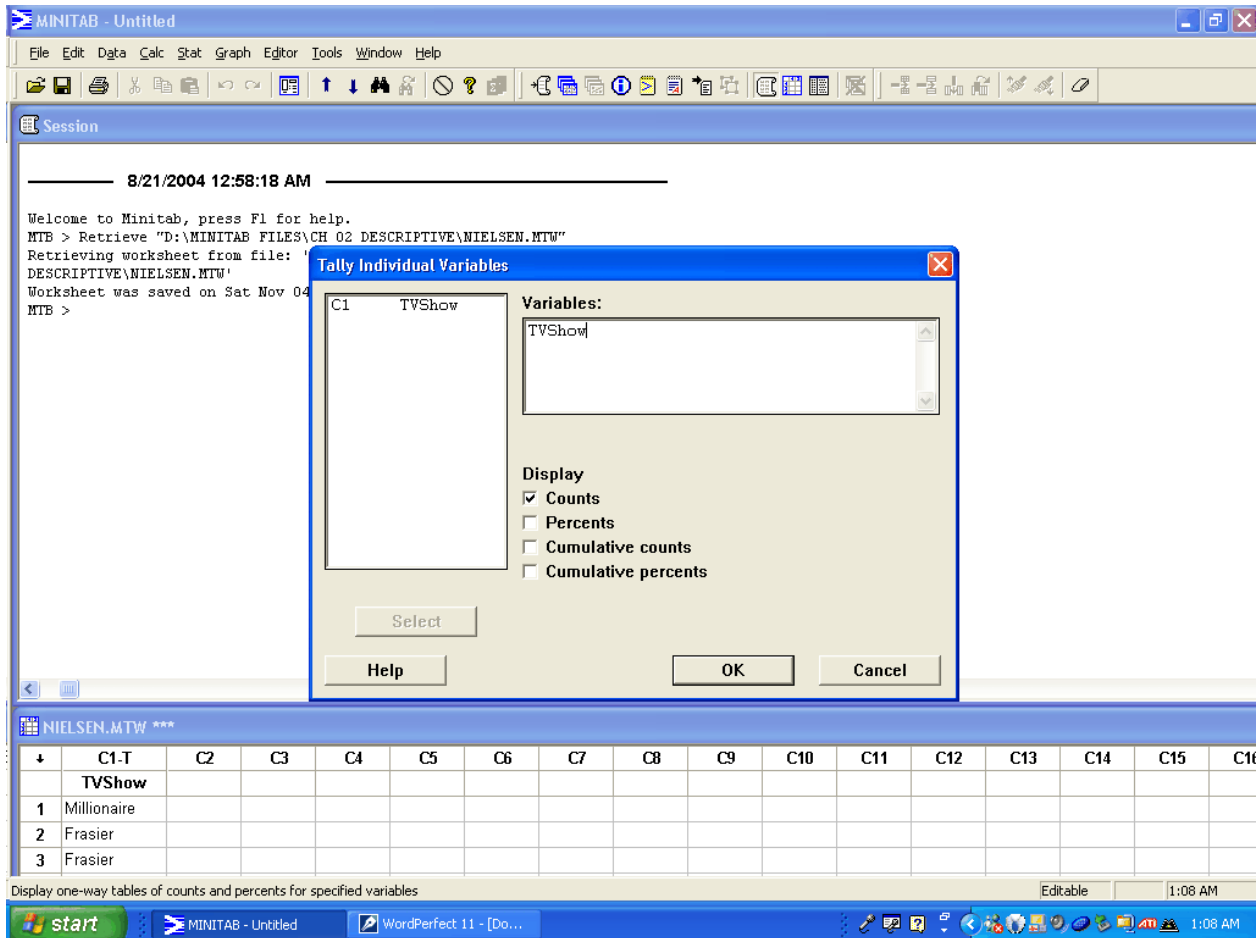


Examples of Graphs for Qualitative Data:

In the *Anderson, Sweeny, and Williams* data sets for Chapter 2, there is a data set showing the show being viewed by 50 viewers in a Nielsen sample. Click on FILE > OPEN WORKSHEET and locate the file “Nielsen” in the chapter two folder.

Click on STAT > TABLES > TALLY INDIVIDUAL VARIABLES

I want to get a count of the number of people watching each of these television shows, so I enter TVShow in the data window and click the “Counts” box.



The screenshot shows the Minitab software interface. A dialog box titled "Tally Individual Variables" is open, showing "TVShow" in the Variables list. Under the "Display" section, the "Counts" checkbox is checked. The background window shows the Minitab session window with a worksheet named "NIELSEN.MTW" containing data for TV shows.

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
	TVShow															
1	Millionaire															
2	Frasier															
3	Frasier															

Click Ok. In the session window you will see this:

Tally for Discrete Variables: TVShow

```
TVShow  Count
Charmed    4
Chicago Hope    7
Frasier    15
Millionaire    24
N=         50
```

Starting with “Charmed” and ending with “24” cut and past these counts into your spreadsheet.

{hint: remove the space between “Chicago” and “Hope” to make it work. Then add it back.}
 Once you have done so you can add some variable names by simply putting the cursor at the head of the column and typing.

The screenshot shows the Minitab software interface. At the top, there is a menu bar (File, Edit, Data, Calc, Stat, Graph, Editor, Tools, Window, Help) and a toolbar. Below the toolbar is a 'Session' window displaying the following text:

```

Millionaire      24
N=              50

MTB >
  
```

Below the session window is a data table window titled 'NIELSEN.MTW ***'. The table has 15 columns labeled C1-T through C15. The first three columns are labeled 'TVShow', 'Show', and 'Viewers' respectively. The data in the table is as follows:

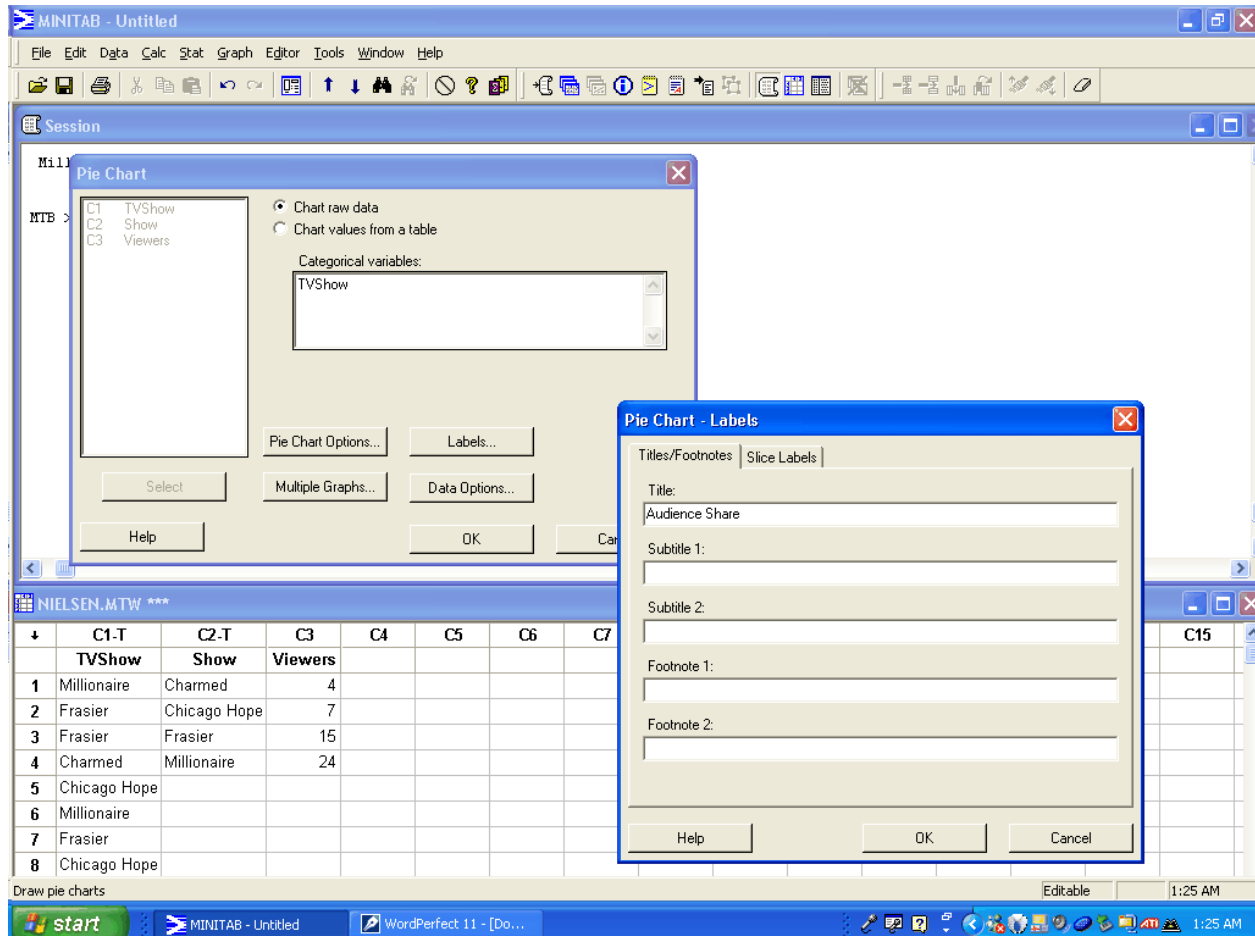
	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	TVShow	Show	Viewers												
1	Millionaire	Charmed	4												
2	Frasier	Chicago Hope	7												
3	Frasier	Frasier	15												
4	Charmed	Millionaire	24												
5	Chicago Hope														
6	Millionaire														
7	Frasier														
8	Chicago Hope														
9	Frasier														
10	Millionaire														
11	Millionaire														
12	Frasier														
13	Millionaire														
14	Millionaire														
15	Charmed														
16	Frasier														
17	Millionaire														
18	Millionaire														
19	Frasier														

At the bottom of the window, there is a status bar with the text 'Welcome to Minitab, press F1 for help.' and the time '1:22 AM'. The Windows taskbar at the very bottom shows the 'start' button and several open applications: 'MINITAB - Untitled', 'WordPerfect 11 - [Do...', and the system clock '1:22 AM'.

Now we can make some pictures.

Click on GRAPH > PIE CHART

We did not have to go to the trouble of using TALLY for this chart. We can proceed as follows, adding a title to the Pie Chart by clicking on the button “Labels.”



We can also ask Minitab to label the pie slices. In the “Labels” box, click on the tab “Slice Labels.”

The screenshot shows the Minitab software interface. The main window displays a data table with columns C1-T through C7. The 'Pie Chart' dialog box is open, showing 'TVShow' as the categorical variable. The 'Pie Chart - Labels' dialog box is also open, with 'Category name', 'Frequency', and 'Percent' selected for labeling the slices.

Session Window:

```

MTB > PieChart
SUBC> Combine
SUBC> Title
SUBC> Panel.

Pie Chart of TV

MTB > PieChart
SUBC> Combine
SUBC> Title
SUBC> SLabel
SUBC> PCate
SUBC> Perce
SUBC> Panel.

Pie Chart of TV

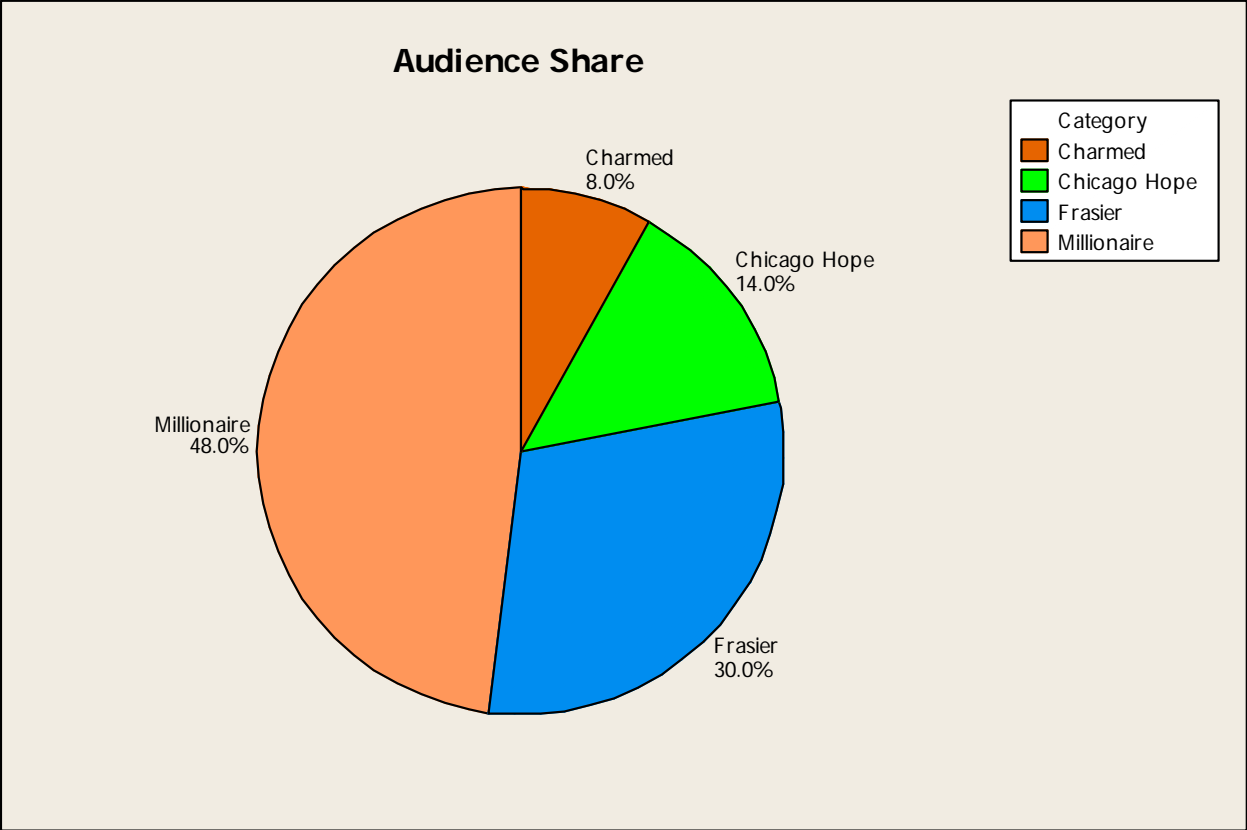
MTB >

```

Table Data:

	C1-T	C2-T	C3	C4	C5	C6	C7
	TVShow	Show	Viewers				
1	Millionaire	Charmed	4				
2	Frasier	Chicago Hope	7				
3	Frasier	Frasier	15				
4	Charmed	Millionaire	24				
5	Chicago Hope						
6	Millionaire						
7	Frasier						
8	Chicago Hope						

Now click on OK > OK, and Minitab generates the following Graph.



Had we originally been given the data as it appears in columns 2 and 3, already tallied, we could have generated the graph in the following way.

The screenshot shows the Minitab software interface. At the bottom, a data table is visible with the following content:

	C1-T	C2-T	C3	C4	C5	C6	C7	C8
	TVShow	Show	Viewers					
1	Millionaire	Charmed	4					
2	Frasier	Chicago Hope	7					
3	Frasier	Frasier	15					
4	Charmed	Millionaire	24					
5	Chicago Hope							
6	Millionaire							
7	Frasier							
8	Chicago Hope							

Two dialog boxes are open over the data table:

- Pie Chart**: This dialog box is set to "Chart values from a table". The "Categorical variable" is "Show" and the "Summary variables" is "Viewers".
- Pie Chart - Labels**: This dialog box is open to the "Titles/Footnotes" tab. The "Title" field contains "Audience Share".

Another thing we can do with this data is make a bar chart. Click on GRAPH > BAR CHART. Select “Counts of unique variable” and “Simple.”

MINITAB - Untitled

File Edit Data Calc Stat Graph Editor Tools Window Help

Session

```

SUBC> Panel.

Pie Chart of TVShow

MTB > PieChart 'TVShow';
SUBC> Combine 0.02;
SUBC> Title "Audience Share";
SUBC> Label;
SUBC> PCategory;
SUBC> Percent;
SUBC> Panel.

Pie Chart of TVShow

MTB > Chart 'TVShow';
SUBC> Bar.

Chart of TVShow

MTB >

```

Bar Charts

Bars represent:

- Counts of unique values
- Counts of unique values
- A function of a variable
- Values from a table

Simple Cluster Stack

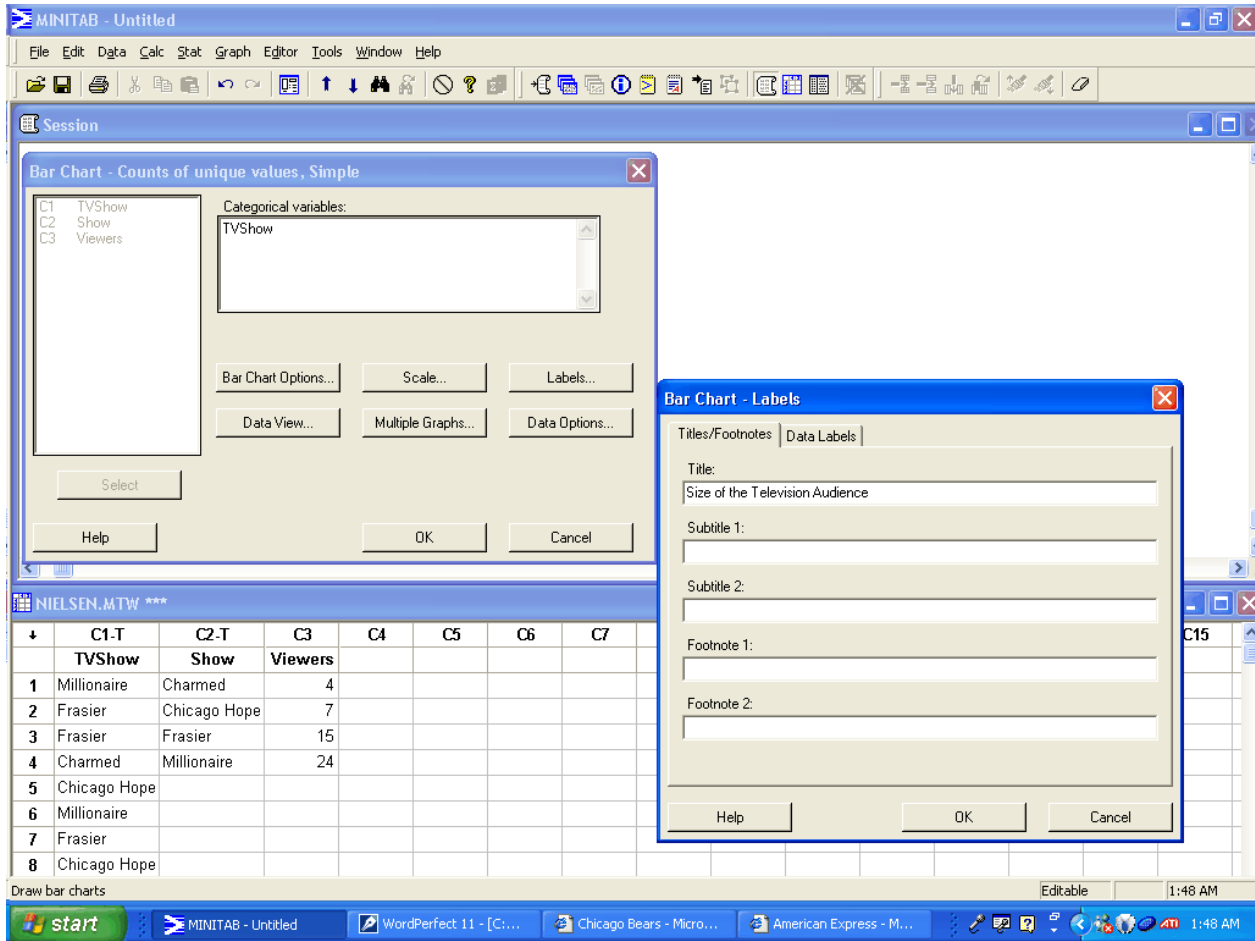
Help OK Cancel

	C1-T	C2-T	C3	C4	C11	C12	C13	C14	C15
	TVShow	Show	Viewers						
1	Millionaire	Charmed	4						
2	Frasier	Chicago Hope	7						
3	Frasier	Frasier	15						
4	Charmed	Millionaire	24						
5	Chicago Hope								
6	Millionaire								
7	Frasier								
8	Chicago Hope								

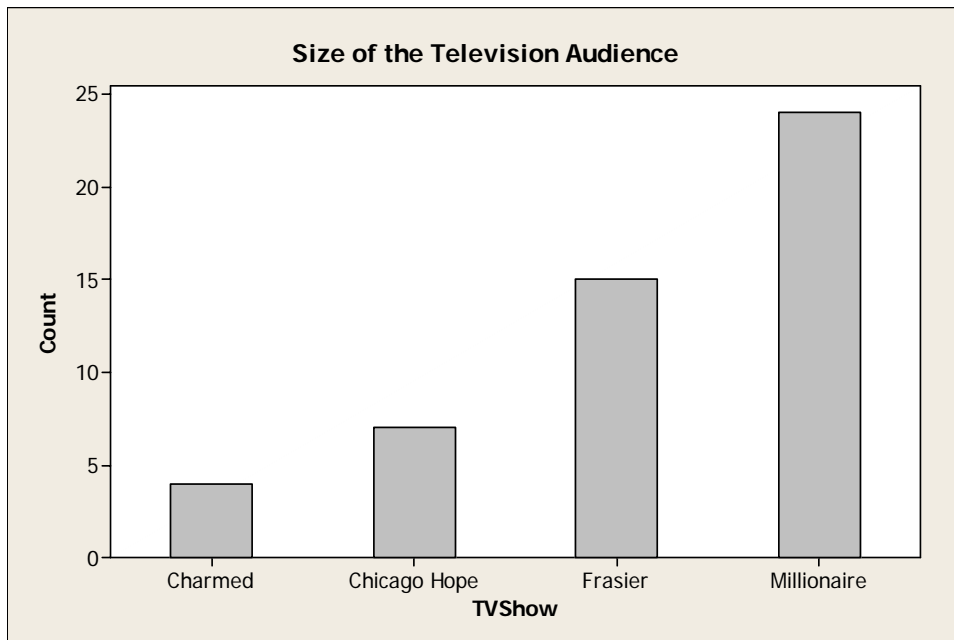
Draw bar charts Editable 1:45 AM

start MINITAB - Untitled WordPerfect 11 - [C:... Chicago Bears - Micro... American Express - M... 1:45 AM

Click on OK. Select your variable and add a label by clicking on the box “Labels.”

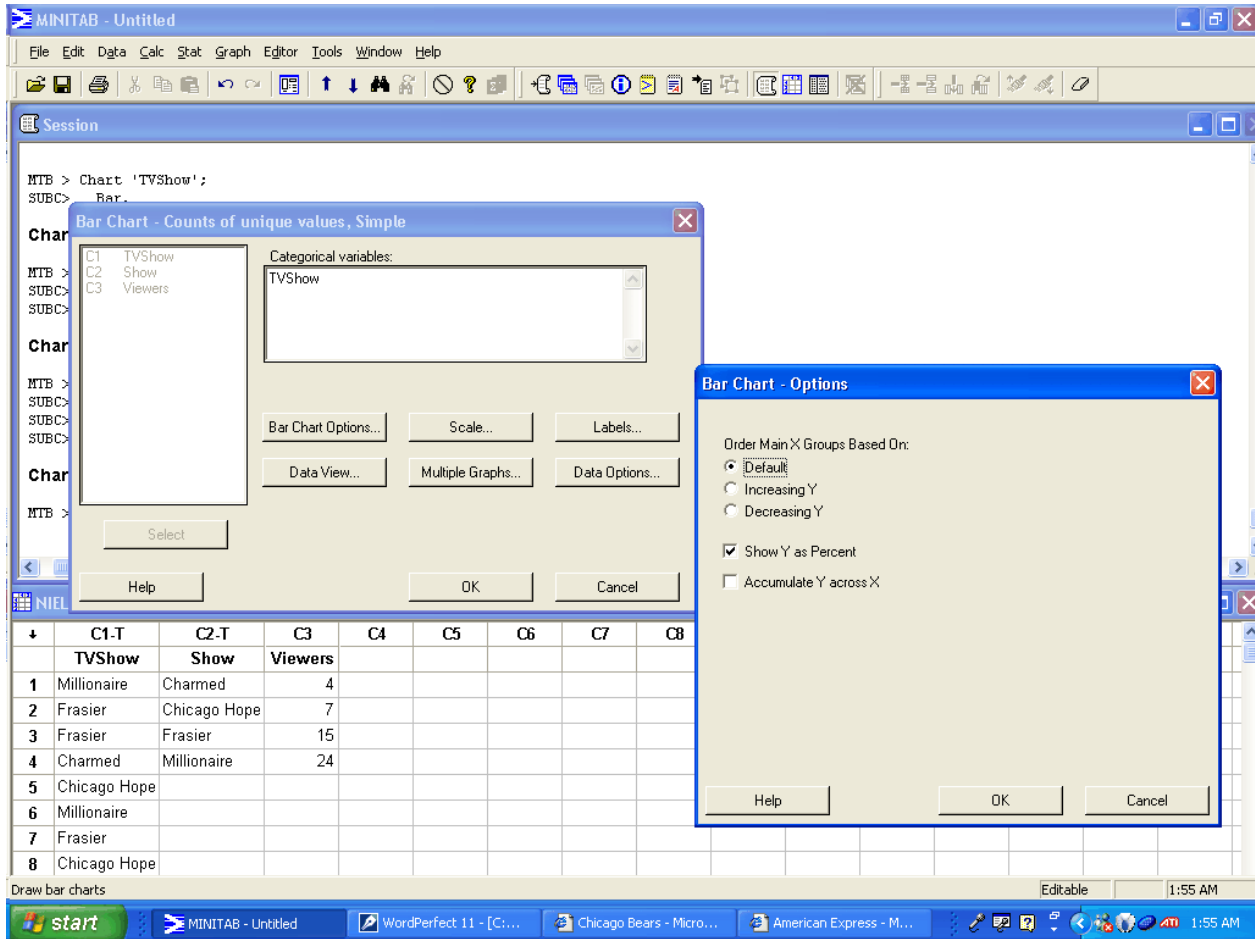


You will get this picture when you click on OK > OK.

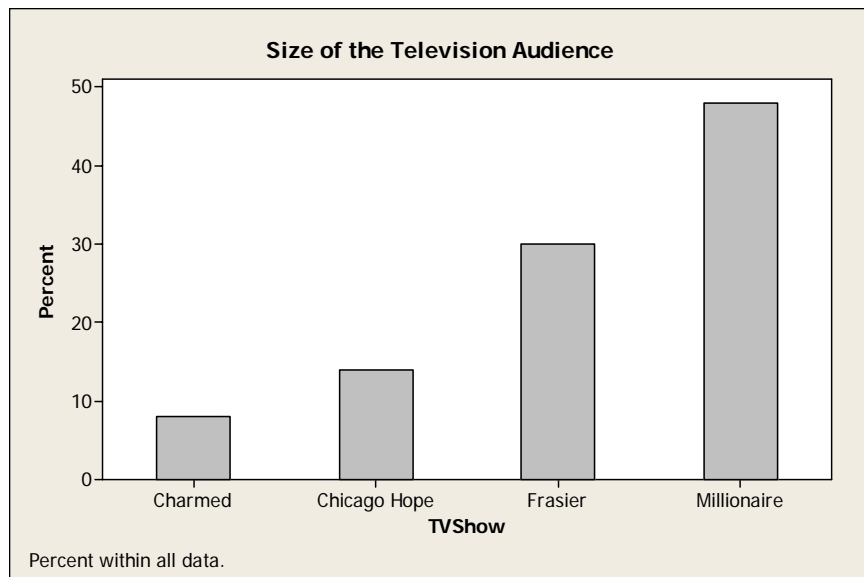


If we

wanted change the Y axis to measure market share in percent, rather than as a count, Minitab will do this for you. Proceed as before, but click on the box “Bar Chart Options.”



Clicking on OK > OK gives this graph.



If the data

had been given to you already tallied, as in columns 2 and 3, we could have made a chart by clicking on BAR CHARTS and then selecting not “Counts of Unique Values,” but “Values from a table,” along with “Simple.” Clicking on OK, we continue thusly:

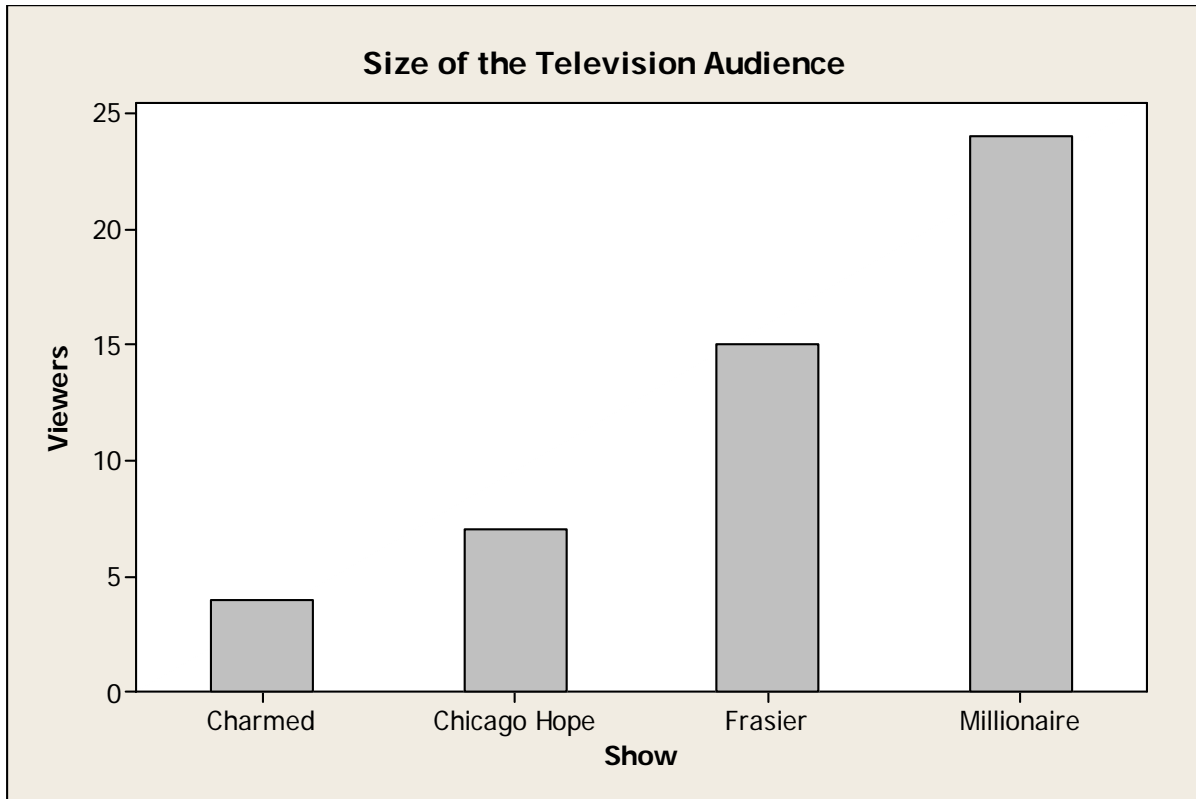
The screenshot shows the Minitab software interface. The main window displays a data table with the following content:

	C1-T	C2-T	C3	C4	C5	C6	C7
	TVShow	Show	Viewers				
1	Millionaire	Charmed	4				
2	Frasier	Chicago Hope	7				
3	Frasier	Frasier	15				
4	Charmed	Millionaire	24				
5	Chicago Hope						
6	Millionaire						
7	Frasier						
8	Chicago Hope						

Two dialog boxes are open over the table:

- Bar Chart - Values from a table, One column of values, Simple:**
 - Graph variables: Viewers
 - Categorical variable: Show
 - Buttons: Bar Chart Options..., Scale..., Labels..., Data View..., Multiple Graphs..., Data Options...
- Bar Chart - Labels:**
 - Titles/Footnotes | Data Labels
 - Title: Size of the Television Audience
 - Subtitle 1: (empty)
 - Subtitle 2: (empty)
 - Footnote 1: (empty)
 - Footnote 2: (empty)
 - Buttons: Help, OK, Cancel

Clicking on OK > OK gives us this graph.



Examples using Quantitative Data

In the Chapter 2 folder of your Anderson, Sweeny, and Williams data disk is a data set, Wageweb, that is a sample of annual salaries (in thousands of dollars) of marketing vice presidents. Open the data set by clicking on FILE > OPEN WORKSHEET and selecting “Wageweb.” We can present this data in many different ways. One way is a dotplot. Click on

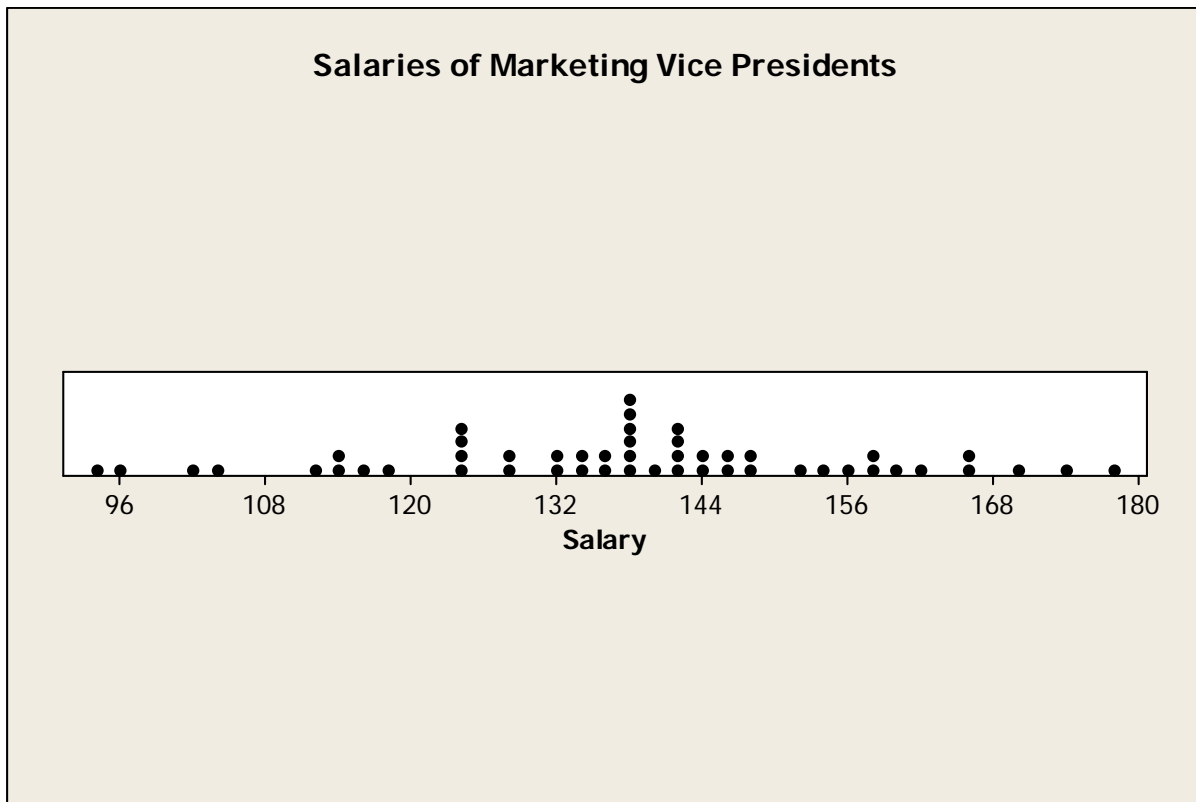
The screenshot shows the Minitab software interface. The main window displays a worksheet titled "Wageweb.MTW ***" with columns C1 through C8. Column C1 is labeled "Salary" and contains the following data points: 145, 140, 145, 173, 116, 127, 155, and 134. Two dialog boxes are open: "Dotplot - One Y, Simple" and "Dotplot - Labels". The "Dotplot - One Y, Simple" dialog box has "C1 Salary" selected in the list of variables and "Salary" entered in the "Graph variables:" field. The "Dotplot - Labels" dialog box has "Titles/Footnotes" selected, and the "Title:" field contains "Salaries of Marketing Vice Presidents".

	C1	C2	C3	C4	C5	C6	C7	C8
	Salary							
1	145							
2	140							
3	145							
4	173							
5	116							
6	127							
7	155							
8	134							

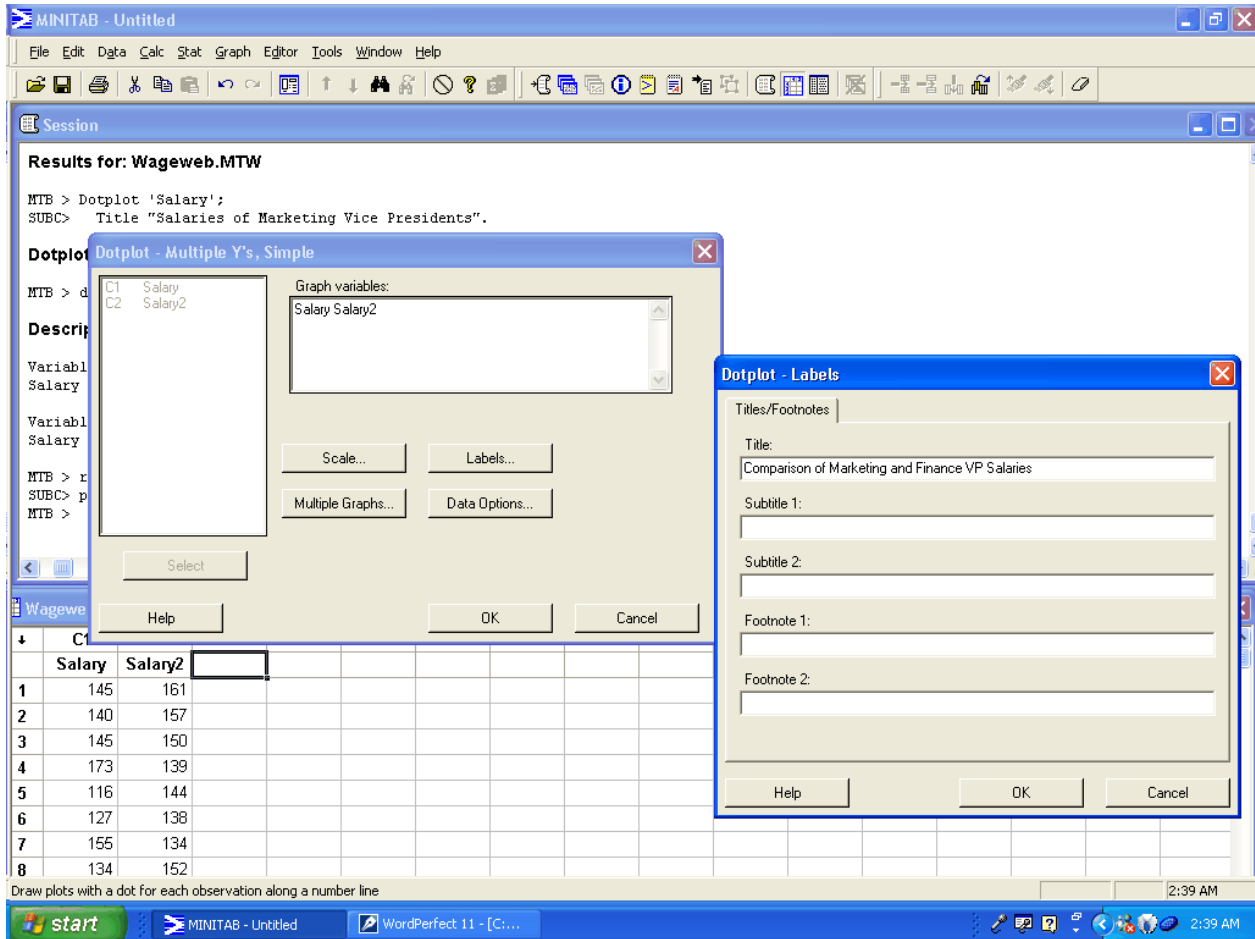
GRAPH > DOTPLOT and select the options “One Y” and “Simple.” Then proceed as follows:

Clicking on OK > OK gives this graph.

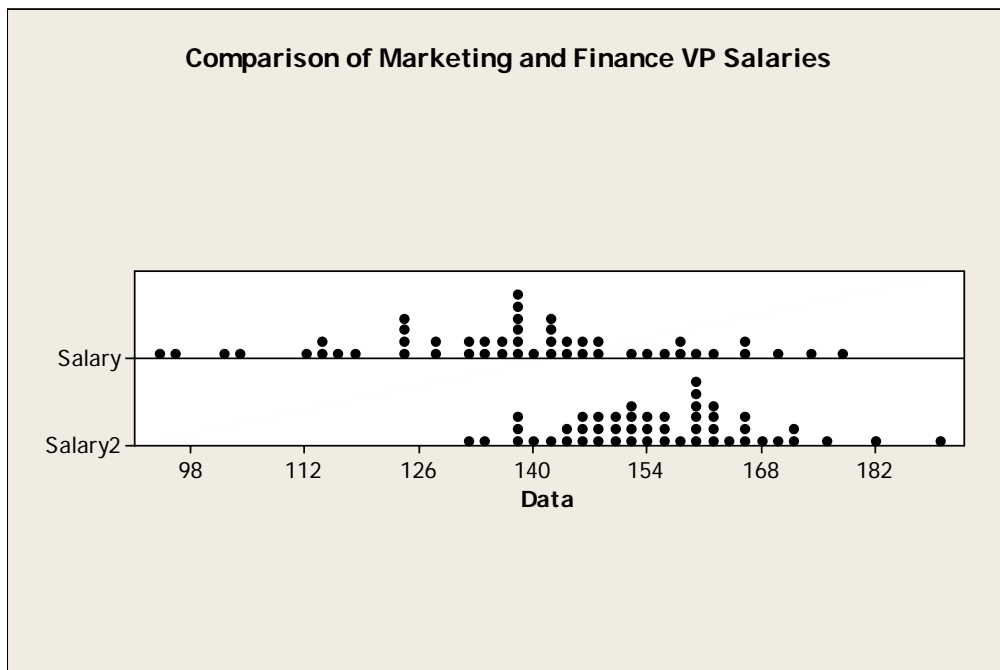
If you had data on a second variable, and wanted a simple visual comparison of the two, ou



could overlay dotplots. For example, suppose you had salary data for Finance Vice Presidents and wanted to compare their pay scale to that of the Marketing Vice Presidents. (I created some fake data, called salary2, to use in this example.) Begin as before, but now select “Multiple Y’s” along with “Simple.” Click OK, and continue in this way:



Click on OK > OK gives this graph.



Another

nice way to look at data is by using a histogram. Click on **GRAPH > HISTOGRAM** and select “Simple.” To make a histogram of Marketing Vice President salaries, do the following:

The screenshot shows the Minitab software interface. The main window displays a data table with the following data:

	C1	C2	C3	C4	C5	C6	C7	C8
	Salary	Salary2						
1	145	161						
2	140	157						
3	145	150						
4	173	139						
5	116	144						
6	127	138						
7	155	134						
8	134	152						

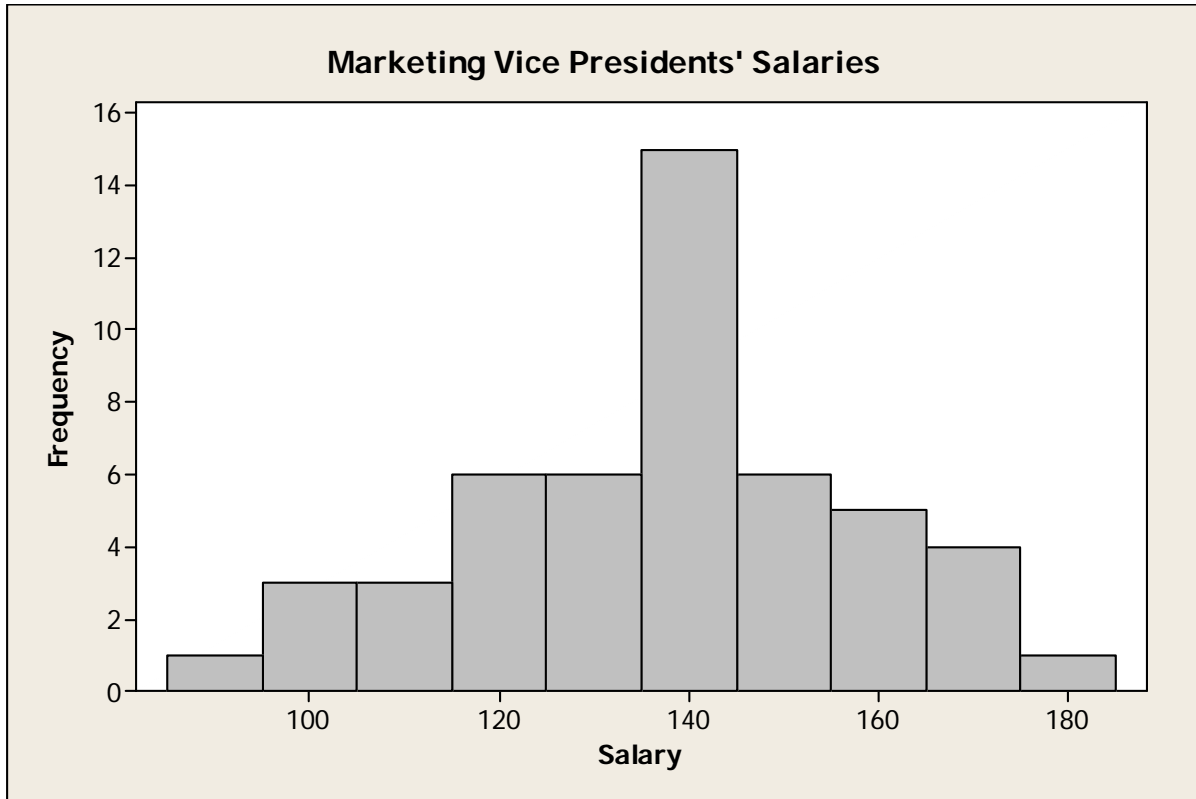
The Histogram - Simple dialog box shows the following configuration:

- Graph variables: Salary
- Buttons: Scale..., Labels..., Data View..., Multiple Graphs..., Data Options..., Select, Help, OK, Cancel

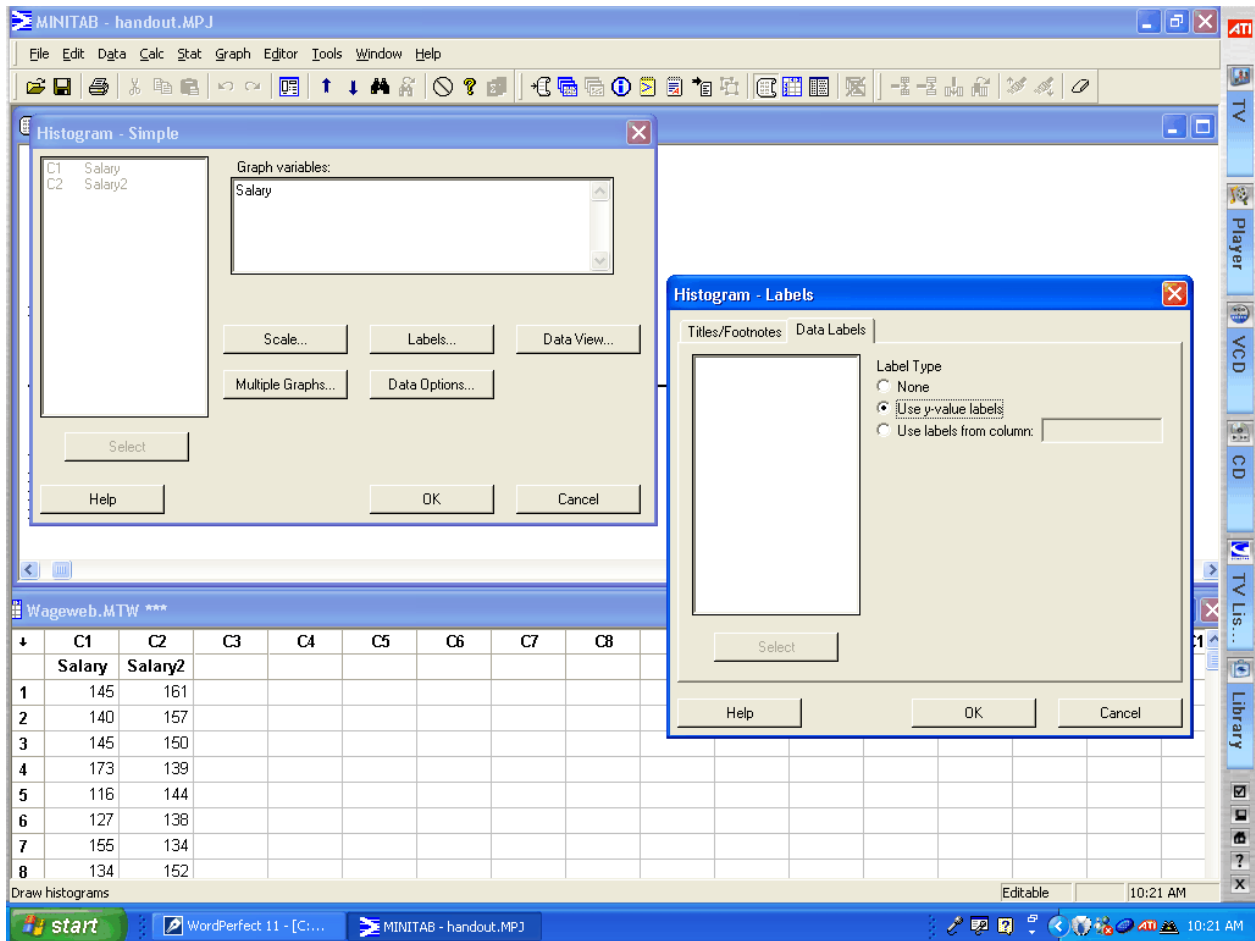
The Histogram - Labels dialog box shows the following configuration:

- Title: Marketing Vice Presidents' Salaries
- Subtitle 1: (empty)
- Subtitle 2: (empty)
- Footnote 1: (empty)
- Footnote 2: (empty)
- Buttons: Help, OK, Cancel

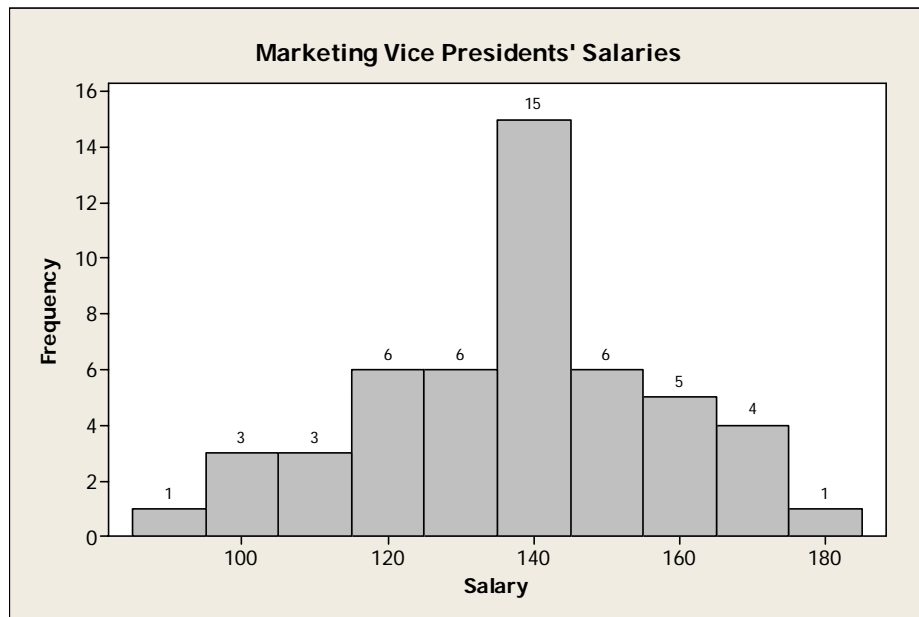
Clicking on OK > OK creates this graph.



There are occasions when it is helpful to display the count in each category. This can be done using one of the options available. Begin by clicking on GRAPH > HISTOGRAM and selecting "Simple." Then click on the "Labels" button and select the tab "Data Labels" and select the radio button "Use Y value labels."

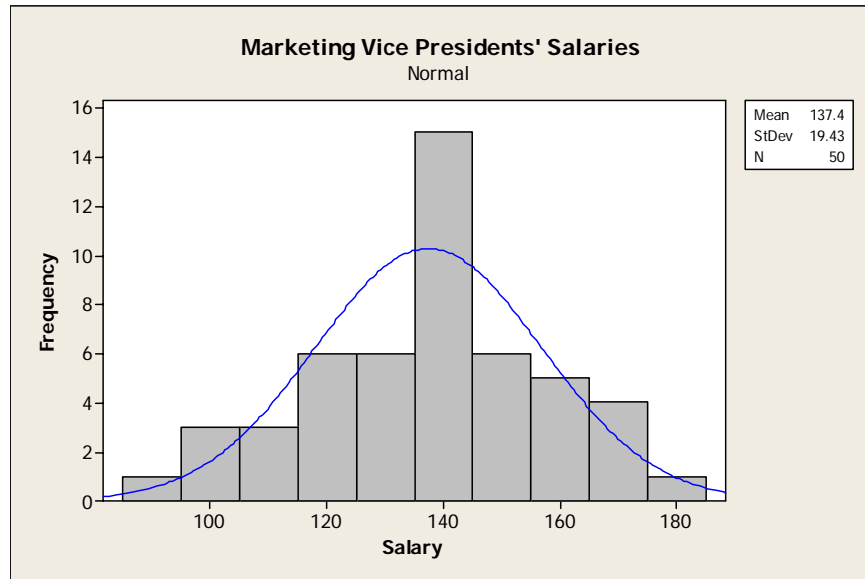


Clicking on OK > OK gives the following graph.

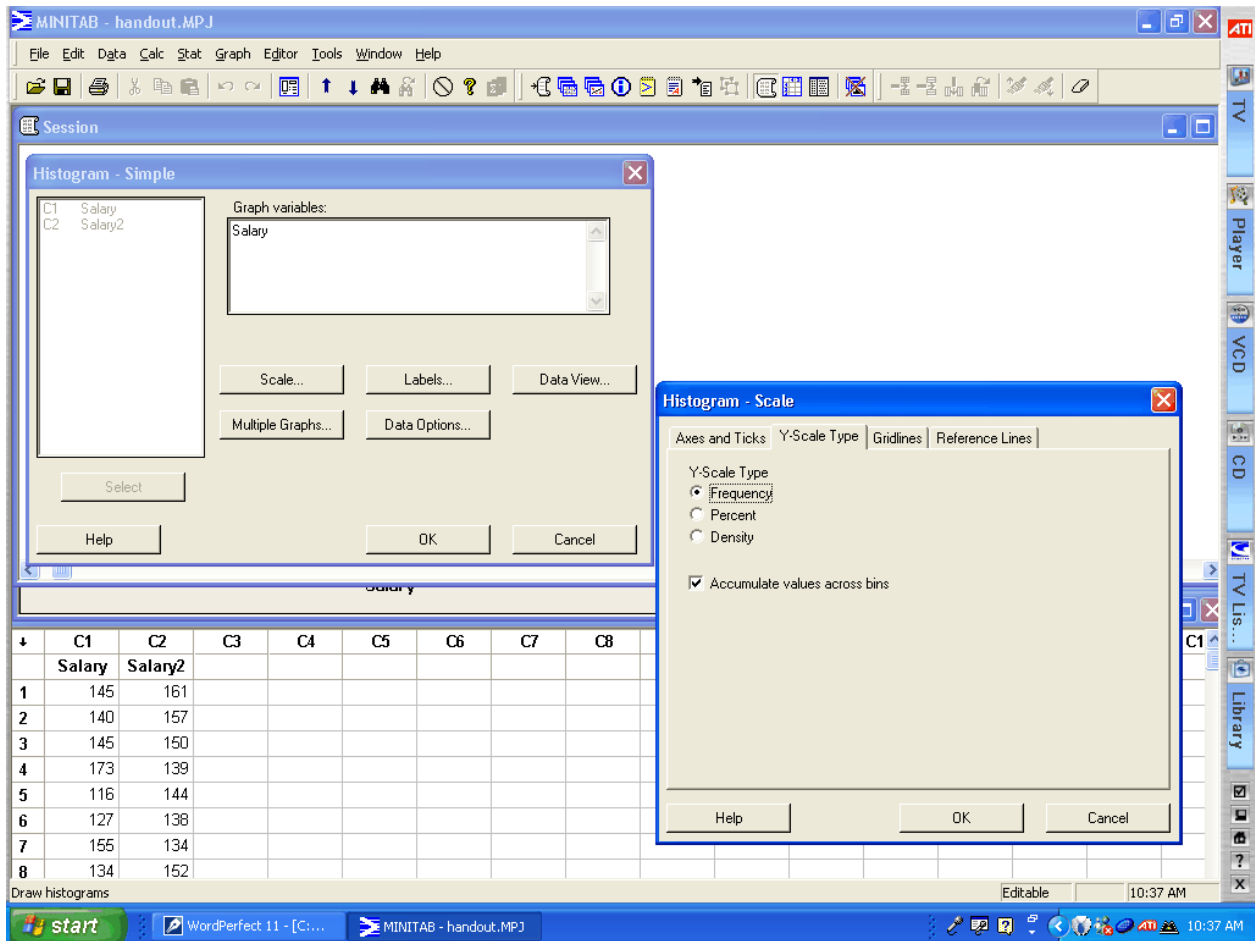


Later in the course we will talk about the normal distribution, the famous “bell curve” people often mention. Sometimes it is useful to see how your histogram compares to the normal distribution, and one way to do so is to create a histogram with an approximating normal distribution superimposed.

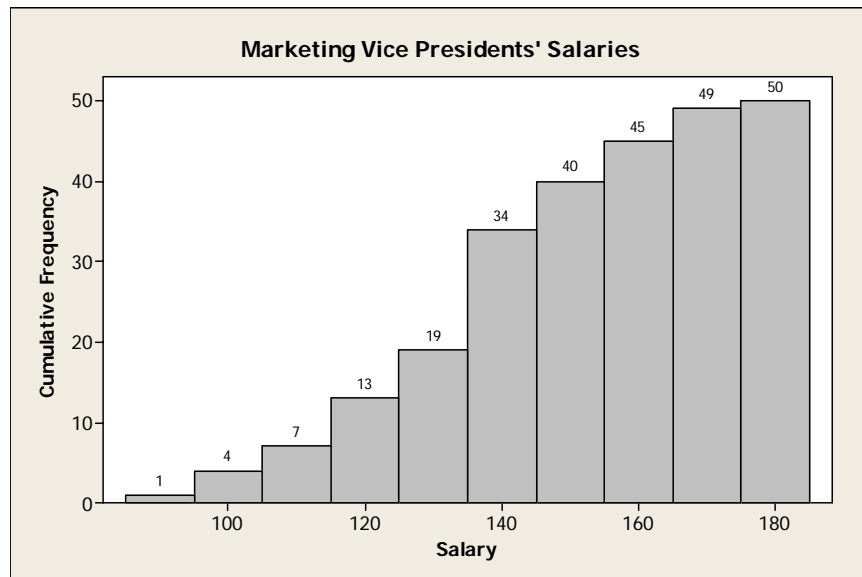
To do so, select GRAPH > HISTOGRAM, but instead of using “Simple,” use “With Fit.” Adding a title, as we did before, yields the following histogram.



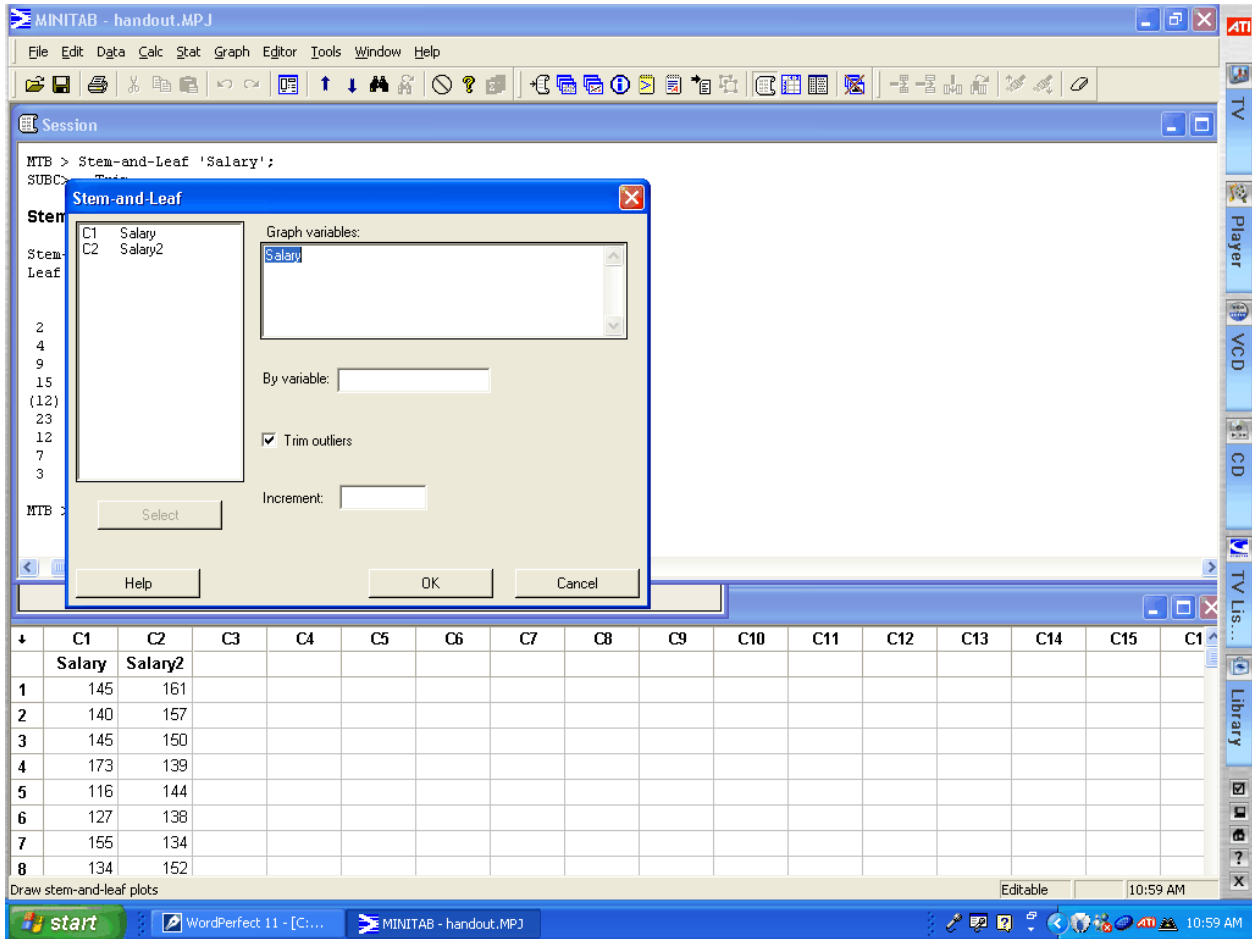
In Anderson, Sweeny, and Williams, pp. 34-36, you will find a discussion of cumulative distributions. These are pictured with a different kind of histogram – one that gives counts that represent the cumulative number *up to* a point. To produce such a histogram in Minitab, begin with GRAPH > HISTOGRAM and select “Simple.” Click on the box “Scale,” and then select the tab called “Y-Scale type.” Check the box “Accumulate values across bins.”



Click on OK > OK, and you get the following histogram.



One of the lesser-used tools discussed in the book is the Stem-and-Leaf display – see pp. 40–43. To create a Stem-and-Leaf display in Minitab, select GRAPH > STEM-AND-LEAF.



When you click OK, you get the following graph, which uncharacteristically for Minitab appears in the Session window.

Stem-and-Leaf Display: Salary

Stem-and-leaf of Salary N = 50
Leaf Unit = 1.0

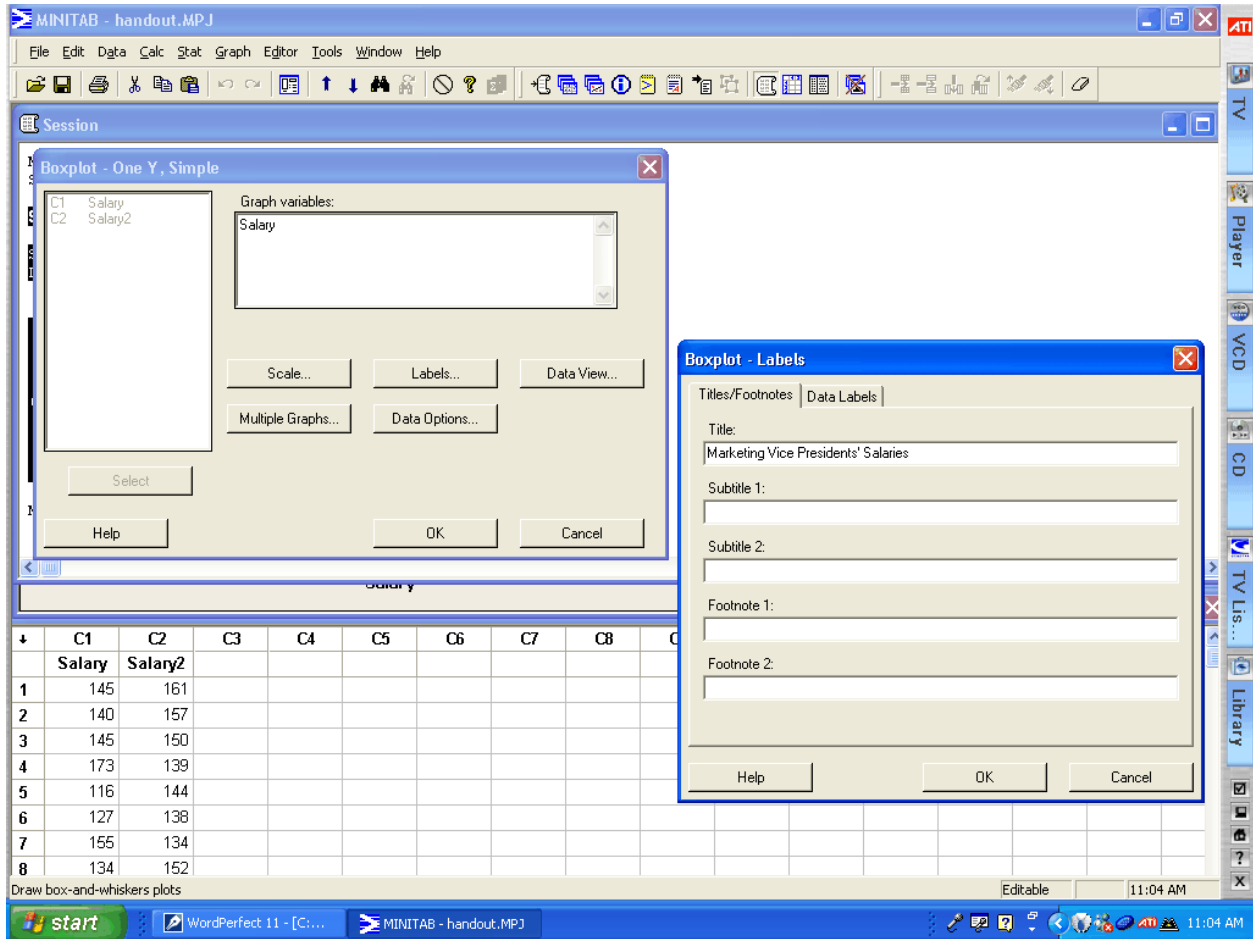
```

2    9    35
4    10   24
9    11   23468
15   12   334477
(12) 13   124456788888
23   14   01122345588
12   15   14577
7    16   0255
3    17   038

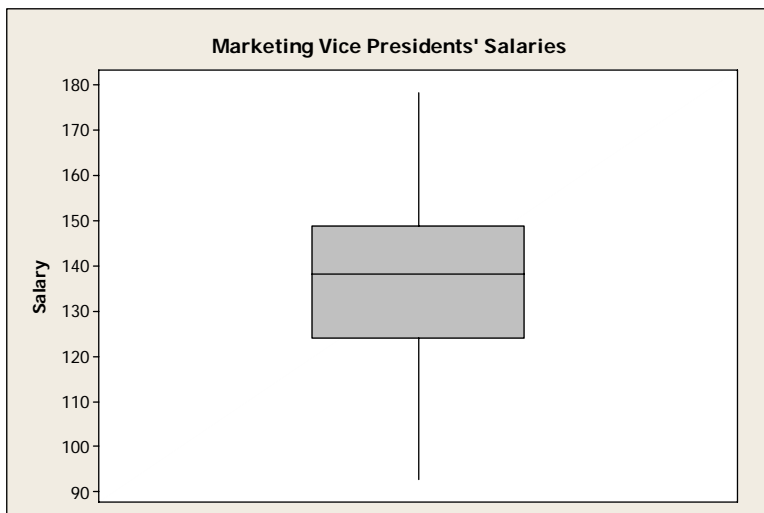
```

There is a final technique for displaying data which is discussed in Anderson, Sweeny, and Williams only in Chapter 3, pp. 101–2, called the Box Plot. To make a box plot of the salary

data in Minitab, click on GRAPH > BOX PLOT, and select “One Y” and “Simple.” Click on “Labels” to add a title.



Clicking on OK > OK gives this boxplot.



Box plots are especially useful for comparing two or more frequency distributions, such as the two salary variables. To display multiple box plots, begin with GRAPH > BOX PLOT, but then select “Multiple Y’s” and “Simple.” For clarity, I am going to rename “Salary” as “Marketing” and “Salary2” as “Finance” by relabelling the head of each column.

The screenshot shows the Minitab software interface. The main window displays a data table with the following content:

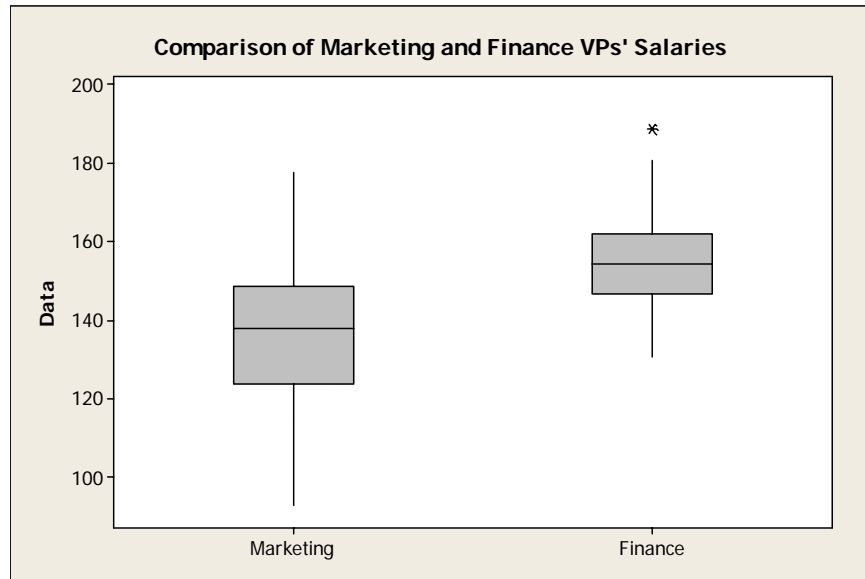
	C1	C2	C3	C4	C5	C6	C7	C8
	Marketing	Finance						
1	145	161						
2	140	157						
3	145	150						
4	173	139						
5	116	144						
6	127	138						
7	155	134						
8	134	152						

Two dialog boxes are open over the main window:

- Boxplot - Multiple Y's, Simple:** This dialog box has a list of variables on the left containing 'C1 Marketing' and 'C2 Finance'. The 'Graph variables:' field contains 'Marketing Finance'. Buttons for 'Scale...', 'Labels...', 'Data View...', 'Multiple Graphs...', and 'Data Options...' are visible.
- Boxplot - Labels:** This dialog box has a 'Titles/Footnotes' tab selected. The 'Title' field contains the text 'Comparison of Marketing and Finance VPs' Salaries'. There are also fields for 'Subtitle 1', 'Subtitle 2', 'Footnote 1', and 'Footnote 2'.

The status bar at the bottom of the Minitab window reads 'Draw box-and-whiskers plots' and the system clock shows '11:12 AM'.

Clicking on OK > OK gives this picture.



Note the asterisk in the Finance box plot, which signifies an outlier.

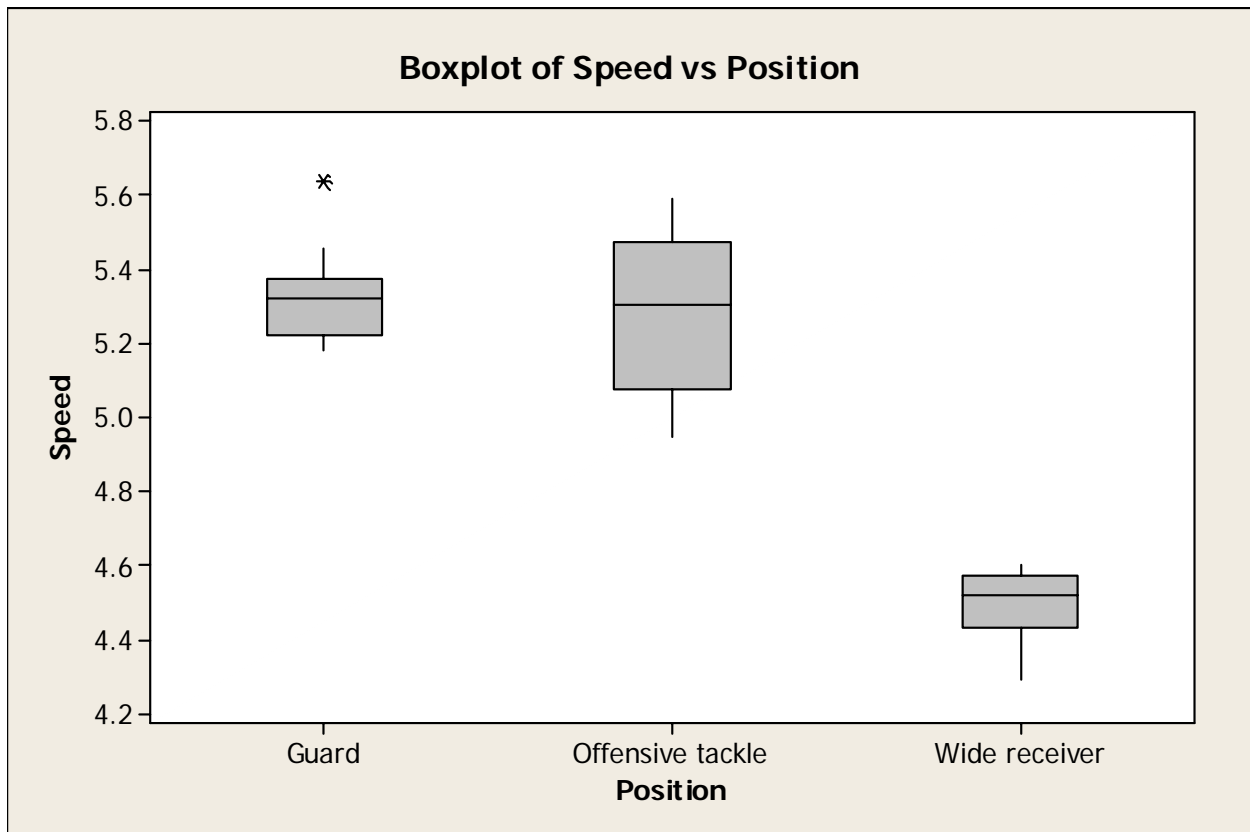
Next, click on FILE > OPEN WORKSHEET and select the file named NFL, which gives information on 40 National football league draft prospects. We can use box plots to compare various attributes of these draft prospects. For example, click on GRAPHS > BOXPLOT and select “One Y” and “With Groups.” This time I propose to accept the default title, and compare prospects’ speed by position.

The screenshot shows the MINITAB software interface. A dialog box titled "Boxplot - One Y, With Groups" is open, allowing the user to configure a boxplot. The "Graph variables:" field contains "Speed", and the "Categorical variables for grouping (1-4, outermost first):" field contains "Position". The background shows a worksheet with the following data:

+	C1	C2-T	C3-T	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Observation	Name	Position	Weight	Speed	Rating								
26	26	Damion McIntosh	Guard	328	5.31	5.3								
27	27	Jeno James	Guard	320	5.64	5.0								
28	28	Al Jackson	Guard	304	5.20	5.0								
29	29	Chris Samuels	Offensive tackle	325	4.95	8.5								
30	30	Stockar McDougle	Offensive tackle	361	5.50	8.0								
31	31	Chris McInosh	Offensive tackle	315	5.39	7.8								
32	32	Adrian Klemm	Offensive tackle	307	4.98	7.6								
33	33	Trid Ward	Offensive tackle	326	5.20	7.3								

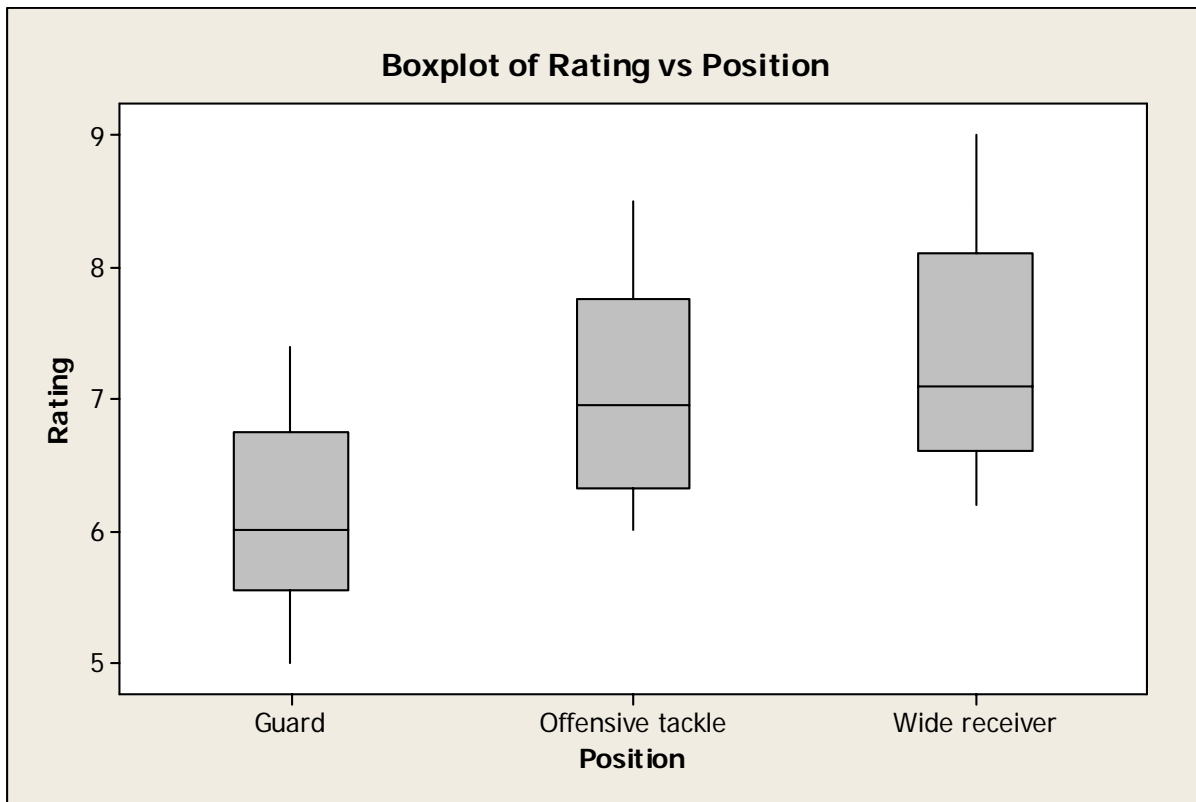
The status bar at the bottom indicates "Draw box-and-whiskers plots" and the time is 11:46 AM.

Clicking on OK gives the following boxplot.



Since “Speed” is time in a 40 yard dash, a low time signifies a speedy individual. What is most apparent (and not at all surprising to any football fan) is that wide receivers are much faster than Guards and Offensive tackles. One less obvious observation is that while the average speed of offensive tackles and guards is almost the same, the speed of offensive tackles is considerably more variable.

Another less obvious result comes from considering the boxplots of rating versus position.



It appears that this particular draft had many blue-chip receiver prospects and few strong prospects at Guard.

Comparing two Qualitative Variables

We can also compare two variables, using a technique known as a Scatterplot. Here is simple example, comparing prospects' speed and weight. Begin by clicking on GRAPH > SCATTERPLOT and selecting "Simple."

MINITAB - handout.MPJ

File Edit Data Calc Stat Graph Editor Tools Window Help

Session

Correlations: Weight, Speed, Rating

Scatterplot - Simple

Y variables: Speed, X variables: Weight

Scatterplot - Labels

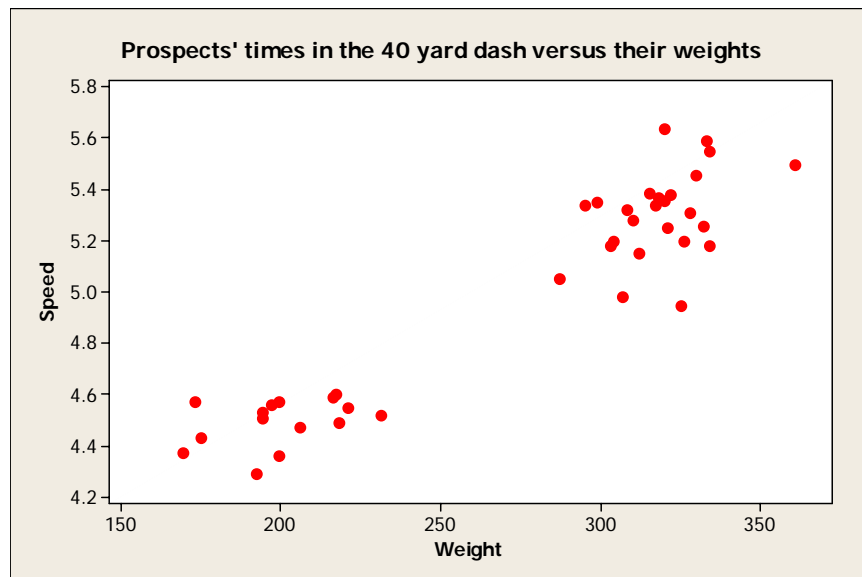
Title: Prospects' times in the 40 yard dash versus their weights

↓	C1	C2-T	C3-T	C4	C5	C6	C7
	Observation	Name	Position	Weight	Speed	Rating	
26	26	Damion McIntosh	Guard	328	5.31	5.3	
27	27	Jeno James	Guard	320	5.64	5.0	
28	28	Al Jackson	Guard	304	5.20	5.0	
29	29	Chris Samuels	Offensive tackle	325	4.95	8.5	
30	30	Stockar McDougle	Offensive tackle	361	5.50	8.0	
31	31	Chris McIngosh	Offensive tackle	315	5.39	7.8	
32	32	Adrian Klemm	Offensive tackle	307	4.98	7.6	
33	33	Todd Wade	Offensive tackle	326	5.20	7.3	

Draw scatterplots Editable 12:11 PM

start WordPerfect 11 - [C:... MINITAB - handout.MPJ 12:11 PM

This produces the following scatterplot.

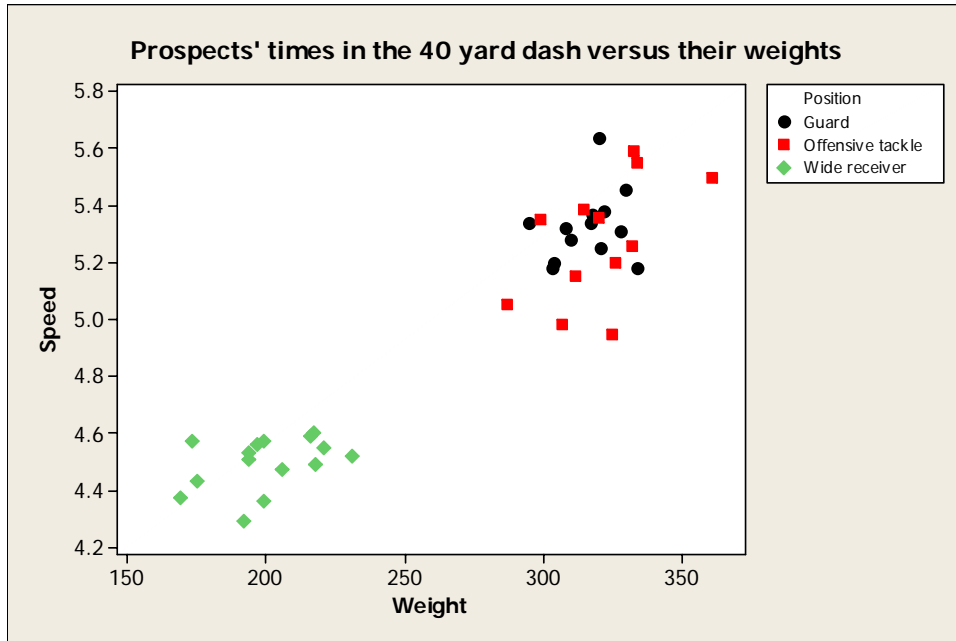


The wide receivers, one might surmise, are the fast and light prospects in the lower left, and the guards and tackles the heavy slow ones in the upper right. We can make this more obvious by going back to GRAPH > SCATTERPLOT and instead of selecting “Simple” selecting “With Groups.”

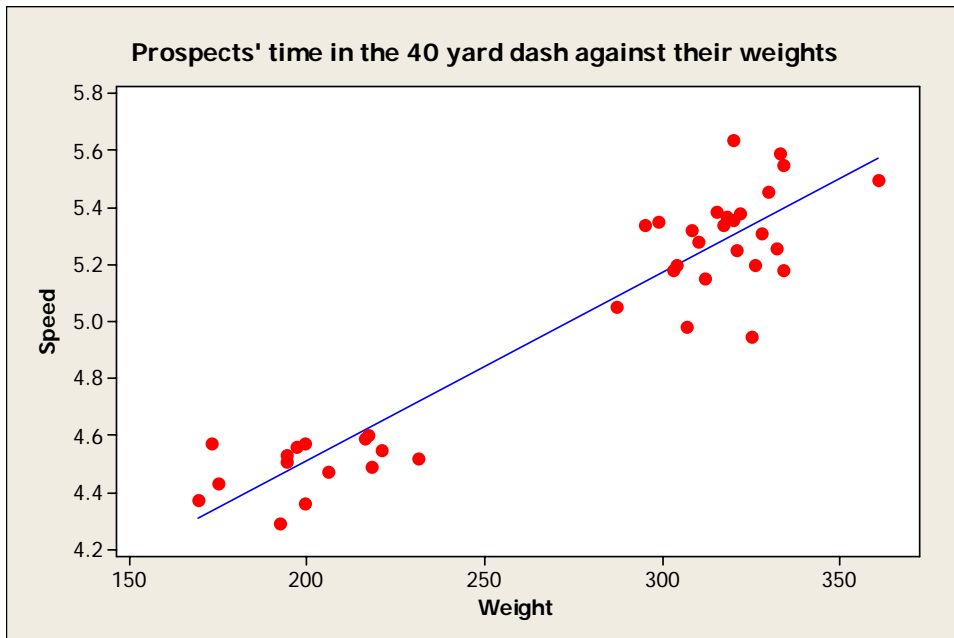
The screenshot shows the Minitab interface with two dialog boxes open. The 'Scatterplot - With Groups' dialog box has 'Speed' selected as the Y variable and 'Weight' as the X variable. 'Position' is selected as the categorical variable for grouping. The 'Scatterplot - Labels' dialog box has the title 'Prospects' times in the 40 yard dash versus their weight'.

↓	C1	C2-T	C3-T	C4	C5	C6	C7
	Observation	Name	Position	Weight	Speed	Rating	
26	26	Damion McIntosh	Guard	328	5.31	5.3	
27	27	Jeno James	Guard	320	5.64	5.0	
28	28	Al Jackson	Guard	304	5.20	5.0	
29	29	Chris Samuels	Offensive tackle	325	4.95	8.5	
30	30	Stockar McDougle	Offensive tackle	361	5.50	8.0	
31	31	Chris McIngosh	Offensive tackle	315	5.39	7.8	
32	32	Adrian Klemm	Offensive tackle	307	4.98	7.6	
33	33	Todd Wade	Offensive tackle	326	5.20	7.3	

This produces the following scatterplot.



We can also produce a fitted line through the scatter of points, if we wish. Begin with GRAPH > SCATTERPLOT and select “With Regression.” Then graph speed against weight.

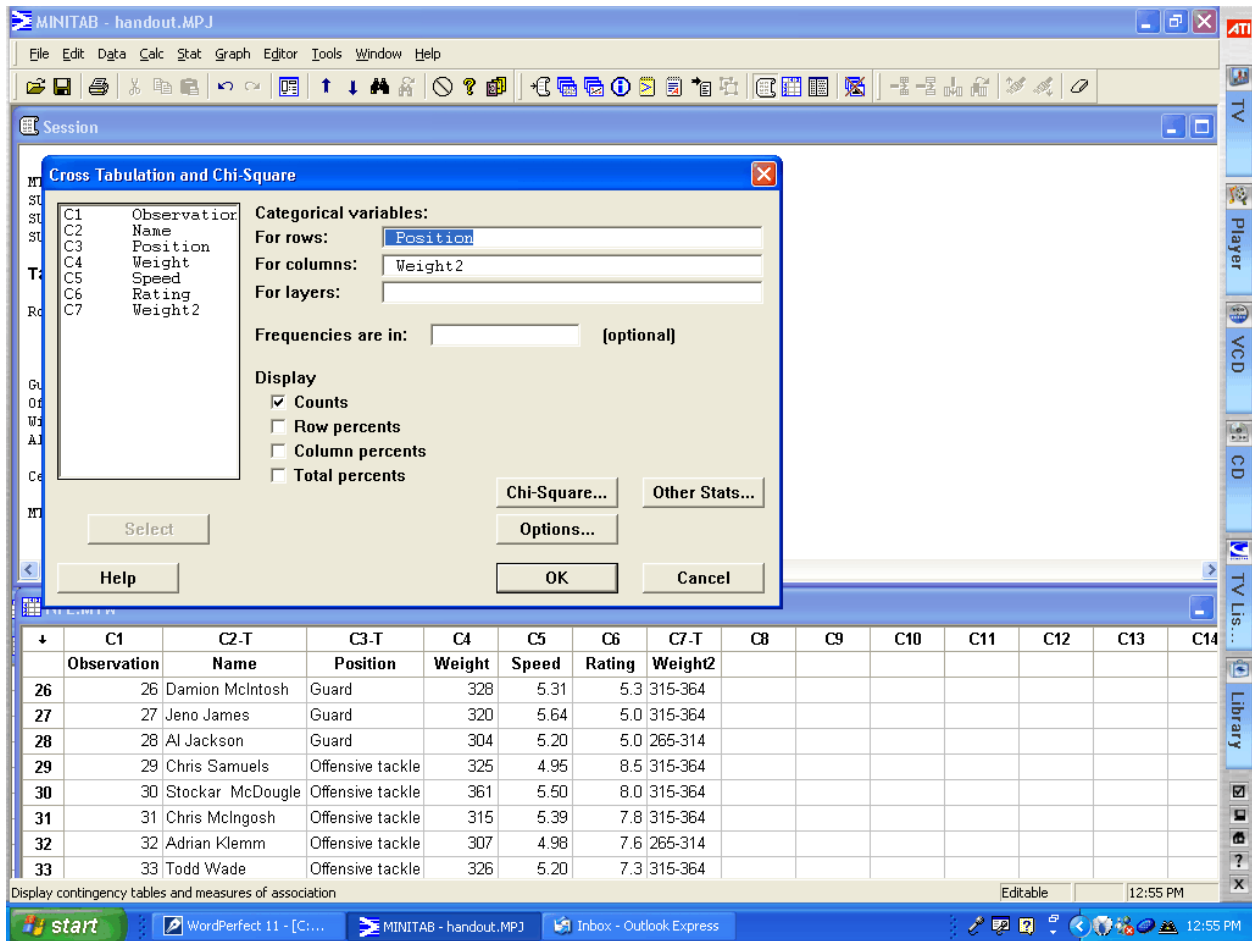


This creates a new qualitative variable, which I have named Weight2. (Qualitative variables have a “T” in the column number. The T is for text.)

The screenshot shows the Minitab interface with a data table for NFL players. The table has the following columns: Observation, Name, Position, Weight, Speed, Rating, and Weight2. The Weight2 column is highlighted in black and contains values ranging from 315-364 to 265-314. The Status bar at the bottom indicates the current worksheet is NFL.MTW.

Observation	Name	Position	Weight	Speed	Rating	Weight2
26	Damion McIntosh	Guard	328	5.31	5.3	315-364
27	Jeno James	Guard	320	5.64	5.0	315-364
28	Al Jackson	Guard	304	5.20	5.0	265-314
29	Chris Samuels	Offensive tackle	325	4.95	8.5	315-364
30	Stockar McDougle	Offensive tackle	361	5.50	8.0	315-364
31	Chris McIngosh	Offensive tackle	315	5.39	7.8	315-364
32	Adrian Klemm	Offensive tackle	307	4.98	7.6	265-314
33	Todd Wade	Offensive tackle	326	5.20	7.3	315-364
34	Marvel Smith	Offensive tackle	320	5.36	7.1	315-364
35	Michael Thompson	Offensive tackle	287	5.05	6.8	265-314
36	Bobby Williams	Offensive tackle	332	5.26	6.8	315-364
37	Darnell Alford	Offensive tackle	334	5.55	6.4	315-364
38	Terrance Beadles	Offensive tackle	312	5.15	6.3	265-314
39	Tutan Reyes	Offensive tackle	299	5.35	6.1	265-314
40	Greg Robinson-Ran	Offensive tackle	333	5.59	6.0	315-364

To Cross-tabulate, click on STAT > TABLES > CROSS TABULATION AND CHI SQUARE.



Clicking on OK results in the following output in the Session window.

Tabulated statistics: Position, Weight2

Rows: Position Columns: Weight2

	165-214	215-264	265-314	315-364	All
Guard	0	0	5	8	13
Offensive tackle	0	0	4	8	12
Wide receiver	10	5	0	0	15
All	10	5	9	16	40