In [1]: **import** pandas **as** pd In [3]: data=pd.read_csv('C:\\Users\\LENOVO\\Downloads\\SaleData.csv') In [4]: type(data) Out[4]: pandas.core.frame.DataFrame In [5]: data.info #concise summary Out[5]: <bound method DataFrame.info of</pre> OrderDate Region Manager SalesMan Units Unit_price \ 1/6/18 East Martha Alexander Television 95.00 1,198.00 1/23/18 Central Hermann Shelli Home Theater 50.00 500.00 2 2/9/18 Central Hermann Luis Television 36.00 1,198.00 Cell Phone 2/26/18 Central Timothy David 27.00 225.00 4 3/15/18 West Timothy Stephen Television 56.00 1,198.00 4/1/18 East Martha Alexander Home Theater 60.00 500.00 Martha Television 4/18/18 Central Steven 75.00 1,198.00 5/5/18 Central Television 90.00 1,198.00 Hermann Luis 5/22/18 8 West Douglas Michael Television 32.00 1,198.00 6/8/18 East Martha Alexander Home Theater 60.00 500.00 10 6/25/18 Central Sigal Television 90.00 1,198.00 Hermann 11 7/12/18 Diana Home Theater 29.00 500.00 East Martha 12 7/29/18 East Douglas Karen Home Theater 81.00 500.00 13 8/15/18 East Martha Alexander Television 35.00 1,198.00 Desk 14 9/1/18 Central Douglas John 2.00 125.00 9/18/18 Alexander Video Games 58.50 15 East Martha 16.00 16 10/5/18 Central Hermann Sigal Home Theater 28.00 500.00 10/22/18 17 East Martha Alexander Cell Phone 64.00 225.00 11/8/18 East Douglas Cell Phone 15.00 225.00 18 Karen 19 11/25/18 Central Hermann Shelli Video Games 96.00 58.50 20 12/12/18 Central Douglas Television John 67.00 1,198.00 East Douglas 21 12/29/18 Karen Video Games 74.00 58.50 1/15/19 Central Timothy David Home Theater 46.00 500.00 22 23 2/1/19 Central John Home Theater 87.00 500.00 Douglas Home Theater 24 2/18/19 East Martha Alexander 4.00 500.00 Stephen Home Theater 25 3/7/19 West Timothy 7.00 500.00 3/24/19 Central 26 Luis Video Games 50.00 58.50 Hermann 27 4/10/19 Central Television 66.00 1,198.00 Martha Steven Cell Phone 28 4/27/19 East Martha Diana 96.00 225.00 29 5/14/19 Central Timothy David Television 53.00 1,198.00 5/31/19 Central David Home Theater 500.00 30 Timothy 80.00 6/17/19 Central Shelli 125.00 31 Hermann Desk 5.00 32 7/4/19 East Martha Alexander Video Games 62.00 58.50 33 7/21/19 Central Hermann Sigal Video Games 55.00 58.50 34 8/7/19 Central Hermann Shelli Video Games 42.00 58.50 35 8/24/19 West Timothy Stephen Desk 3.00 125.00 9/10/19 Central 36 Timothy David Television 7.00 1,198.00 9/27/19 Cell Phone 37 West Timothy Stephen 76.00 225.00 38 10/14/19 Michael Home Theater 57.00 500.00 West Douglas 39 10/31/19 Central Martha Steven Television 14.00 1,198.00 11/17/19 40 Central Luis Home Theater 11.00 500.00 Hermann 41 12/4/19 Central Hermann Luis Home Theater 94.00 500.00 42 12/21/19 Central Steven Home Theater 28.00 500.00 Martha 43 NaN NaN 278.00 1,125.00 NaN NaN NaN 44 NaN NaN NaN NaN NaN 34.75 140.63 Sale_amt 113,810.00 0 25,000.00 43,128.00 6,075.00 3 67,088.00 30,000.00 89,850.00 107,820.00 38,336.00 9 30,000.00 107,820.00 10 11 14,500.00 12 40,500.00 13 41,930.00 14 250.00 15 936.00 14,000.00 16 17 14,400.00 18 3,375.00 19 5,616.00 20 80,266.00 21 4,329.00 22 23,000.00 23 43,500.00 24 2,000.00 25 3,500.00 26 2,925.00 27 79,068.00 28 21,600.00 29 63,494.00 40,000.00 30 31 625.00 32 3,627.00 33 3,217.50 34 2,457.00 35 375.00 36 8,386.00 37 17,100.00 28,500.00 39 16,772.00 40 5,500.00 41 47,000.00 42 14,000.00 43 62,550.00 7,818.75 > In [6]: data.describe() #descriptive statistics Out[6]: Units count 45.000000 54.083333 mean 45.096676 2.000000 min 25% 28.000000 53.000000 50% 75.000000 75% max 278.000000 In [8]: data=data.drop_duplicates() Out[8]: OrderDate Region Manager SalesMan Units Unit_price Sale_amt Item 0 1/6/18 East Martha Alexander Television 95.00 1,198.00 113,810.00 1/23/18 Central Hermann Shelli Home Theater 50.00 500.00 25,000.00 2/9/18 Central Hermann 1,198.00 43,128.00 2 Television 36.00 Luis 2/26/18 Central 225.00 6,075.00 David Cell Phone 27.00 Timothy Stephen 56.00 1,198.00 67,088.00 3/15/18 West Timothy Television 4/1/18 Martha Alexander Home Theater 60.00 500.00 30,000.00 East 4/18/18 Central Martha Steven Television 75.00 1,198.00 89,850.00 5/5/18 Central Hermann 90.00 1,198.00 107,820.00 Luis Television 5/22/18 West Douglas Michael Television 32.00 1,198.00 38,336.00 East Martha Alexander Home Theater 60.00 6/8/18 500.00 30,000.00 10 6/25/18 Central Hermann Television 90.00 1,198.00 107,820.00 Sigal 500.00 14,500.00 11 7/12/18 East Martha Diana Home Theater 29.00 12 7/29/18 East Douglas Karen Home Theater 81.00 500.00 40,500.00 1,198.00 41,930.00 8/15/18 Television 35.00 13 East Martha Alexander 9/1/18 Central 14 Douglas John Desk 2.00 125.00 250.00 9/18/18 Martha Alexander Video Games 16.00 58.50 936.00 15 East 16 10/5/18 Central Hermann Sigal Home Theater 28.00 500.00 14,000.00 Cell Phone 64.00 225.00 14,400.00 17 10/22/18 East Martha Alexander 18 11/8/18 East Douglas Cell Phone 15.00 225.00 3,375.00 Karen 11/25/18 Central Hermann Video Games 96.00 58.50 5,616.00 Shelli 12/12/18 Central Douglas John Television 67.00 1,198.00 80,266.00 Karen Video Games 74.00 4,329.00 **21** 12/29/18 East Douglas 58.50 1/15/19 Central David Home Theater Timothy 46.00 500.00 23,000.00 Douglas 43,500.00 23 2/1/19 Central 87.00 500.00 John Home Theater 24 2/18/19 East Martha Alexander Home Theater 4.00 500.00 2,000.00 3/7/19 Stephen Home Theater 7.00 500.00 3,500.00 25 West Timothy 26 3/24/19 Central Hermann Luis Video Games 50.00 58.50 2,925.00 1,198.00 79,068.00 4/10/19 Central 66.00 27 Martha Steven Television 28 4/27/19 East Martha Diana Cell Phone 96.00 225.00 21,600.00 5/14/19 Central Timothy 1,198.00 63,494.00 Television 53.00 29 David 30 5/31/19 Central Timothy David Home Theater 80.00 500.00 40,000.00 31 6/17/19 Central Hermann Shelli Desk 5.00 125.00 625.00 32 Martha Alexander Video Games 62.00 58.50 3,627.00 East 7/21/19 Central Hermann 55.00 3,217.50 33 Video Games 58.50 34 8/7/19 Central Hermann Shelli Video Games 42.00 58.50 2,457.00 8/24/19 Stephen 3.00 125.00 375.00 West Timothy Desk 9/10/19 Central Timothy David Television 7.00 1,198.00 8,386.00 9/27/19 West Timothy Stephen Cell Phone 76.00 225.00 17,100.00 37 10/14/19 West Douglas 28,500.00 Michael Home Theater 500.00 10/31/19 Central 14.00 1,198.00 16,772.00 Martha Television 11/17/19 Central Hermann Luis Home Theater 11.00 500.00 5,500.00 12/4/19 Central Hermann Luis Home Theater 94.00 500.00 47,000.00 Steven Home Theater 28.00 42 12/21/19 Central 500.00 14,000.00 Martha 62,550.00 NaN NaN NaN 278.00 1,125.00 NaN NaN 44 NaN NaN NaN 34.75 140.63 7,818.75 NaN NaN In [9]: data.isnull() Out[9]: OrderDate Region Manager SalesMan Item Units Unit_price Sale_amt 0 False 2 False 4 False 8 False False False False False False False False 9 False 10 False False False False False False False False 11 False False False False False 12 False False False False False False False False 13 False False False False False False False False 14 False 15 False False False False False 16 False False False False False False False 17 False False False False False False False False 18 False 19 False False False False False 20 False 21 False False False False 22 False False False False False False False False 23 False False False False False False False False 24 False 26 False False False False False False 27 False False False False False False False False 28 False 29 False False False False 30 False False False False False False False False False 31 False False False False False False False 32 False False False False False False False False 33 False False False False False False False False 34 False False False False False False False False 35 False False False False False False False False 36 False 37 False False False 38 False 39 False False False False 40 False False False False False False False False 41 False False False False False False False False 42 False False False False False False False 43 True True True True True False False False 44 True True False True True True False False In [10]: data.isnull().sum() Out[10]: OrderDate 2 Region 2 Manager 2 SalesMan 2 Item 2 Units 0 Unit_price 0 Sale_amt dtype: int64 In [11]: data.isnull().sum().sum() Out[11]: **10** In [12]: data.notnull() Out[12]: OrderDate Region Manager SalesMan Item Units Unit_price Sale_amt True 8 True 10 True 11 True True 12 True True True True True True True True 13 True 14 True 15 True 16 True True True True True 17 True True True True True True True True 18 True True True True True True True True 19 True 20 True True True True 21 True 22 True True True True True True 23 True 24 True True True True True 25 True True True True True True True True 26 True True True True True True True True 27 True True True True True True True True 28 True 29 True True True 30 True True True True True 31 True 32 True True True True True 33 True 34 True 35 True True 36 True True True True True True True True 37 True 38 True 42 True True True True True True True True False False 43 False False False True True 44 False False False False False True True True In [13]: data2=data.fillna(value=0) Out[13]: OrderDate Region Manager SalesMan Item Units Unit_price Sale_amt 1/6/18 East Martha Alexander Television 1,198.00 113,810.00 1/23/18 Central Hermann 50.00 500.00 25,000.00 Shelli Home Theater 43,128.00 2/9/18 Central Hermann Television 36.00 1,198.00 2/26/18 Central David Cell Phone 27.00 225.00 6,075.00 Timothy 56.00 1,198.00 67,088.00 3/15/18 West Timothy Stephen Television Martha Alexander Home Theater 60.00 4/1/18 East 500.00 30,000.00 4/18/18 Central Martha Steven Television 75.00 1,198.00 89,850.00 90.00 1,198.00 107,820.00 5/5/18 Central Hermann Television Luis 1,198.00 38,336.00 5/22/18 West Douglas Michael Television 32.00 6/8/18 Martha Alexander Home Theater 60.00 500.00 30,000.00 East 10 6/25/18 Central Hermann Television 90.00 1,198.00 107,820.00 7/12/18 29.00 500.00 14,500.00 11 East Martha Diana Home Theater 81.00 500.00 40,500.00 12 7/29/18 East Douglas Karen Home Theater 8/15/18 35.00 1,198.00 41,930.00 13 East Martha Alexander Television 14 9/1/18 Central Douglas Desk 2.00 125.00 250.00 9/18/18 East Martha Alexander Video Games 16.00 58.50 936.00 500.00 14,000.00 16 10/5/18 Central Hermann Sigal Home Theater 28.00 **17** 10/22/18 Martha Alexander Cell Phone 64.00 225.00 14,400.00 East 3,375.00 18 11/8/18 East Douglas Karen Cell Phone 15.00 225.00 19 11/25/18 Central Hermann Video Games 96.00 58.50 5,616.00 Shelli 12/12/18 Central Douglas John Television 67.00 1,198.00 80,266.00 Video Games 74.00 **21** 12/29/18 Douglas 58.50 4,329.00 East Karen David Home Theater 22 1/15/19 Central Timothy 46.00 500.00 23,000.00 2/1/19 Central John Home Theater 87.00 500.00 43,500.00 23 Douglas 500.00 2,000.00 24 2/18/19 East Martha Alexander Home Theater 4.00 25 West Timothy Stephen Home Theater 7.00 500.00 3,500.00 3/24/19 Central Hermann Luis Video Games 50.00 58.50 2,925.00 **27** 4/10/19 Central Martha Steven Television 66.00 1,198.00 79,068.00 East Martha Cell Phone 96.00 225.00 21,600.00 5/14/19 Central Television 53.00 1,198.00 63,494.00 29 Timothy David 30 5/31/19 Central Timothy David Home Theater 80.00 500.00 40,000.00 6/17/19 Central Hermann 5.00 125.00 625.00 31 Shelli Desk 3,627.00 32 East Martha Alexander Video Games 62.00 58.50 7/21/19 Central Hermann Video Games 55.00 58.50 3,217.50 33 Sigal 34 8/7/19 Central Hermann Shelli Video Games 42.00 58.50 2,457.00 35 8/24/19 West Stephen Desk 3.00 125.00 375.00 Timothy 9/10/19 Central 8,386.00 Timothy David Television 7.00 1,198.00 Cell Phone 76.00 225.00 17,100.00 37 9/27/19 West Stephen Timothy 500.00 West Douglas Michael Home Theater 57.00 28,500.00 10/31/19 Central 14.00 1,198.00 16,772.00 Martha Steven Television 11/17/19 Central Hermann Luis Home Theater 11.00 500.00 5,500.00 12/4/19 Central Hermann Luis Home Theater 94.00 500.00 47,000.00 12/21/19 Central Martha Steven Home Theater 28.00 500.00 14,000.00 43 0 0 0 278.00 1,125.00 62,550.00 44 0 0 34.75 140.63 7,818.75 In [14]: data3=data.fillna(value='pad') data3 Out[14]: OrderDate Region Manager SalesMan Units Unit_price Sale_amt 1,198.00 113,810.00 1/6/18 East Martha Alexander Television 95.00 1/23/18 Central Hermann 50.00 500.00 25,000.00 Shelli Home Theater 2 2/9/18 Central Hermann Television 36.00 1,198.00 43,128.00 Cell Phone 27.00 225.00 6,075.00 2/26/18 Central Timothy David 3/15/18 Stephen 56.00 1,198.00 67,088.00 West Timothy Television Martha Alexander Home Theater 60.00 500.00 30,000.00 East 1,198.00 89,850.00 4/18/18 Central Television 75.00 Martha Steven 1.198.00 107.820.00 90.00 5/5/18 Central Hermann Luis Television 5/22/18 Michael Television 32.00 1,198.00 38,336.00 Douglas 6/8/18 East Martha Alexander Home Theater 60.00 500.00 30,000.00 10 6/25/18 Central Hermann Television 90.00 1,198.00 107,820.00 500.00 14,500.00 7/12/18 Diana Home Theater 29.00 11 East Martha 500.00 40,500.00 12 7/29/18 Douglas Karen Home Theater 81.00 East 8/15/18 Television 35.00 1,198.00 41,930.00 East Martha Alexander 14 9/1/18 Central Douglas John Desk 2.00 125.00 Martha Alexander Video Games 16.00 9/18/18 58.50 936.00 10/5/18 Home Theater 28.00 500.00 14,000.00 10/22/18 Cell Phone 64.00 225.00 14,400.00 17 East Martha Alexander 18 11/8/18 Cell Phone 15.00 225.00 3,375.00 East Douglas Karen 11/25/18 Central Hermann Shelli Video Games 96.00 58.50 5,616.00 12/12/18 Central 67.00 80,266.00 20 Douglas John Television 1,198.00 **21** 12/29/18 Karen Video Games 74.00 58.50 4,329.00 East Douglas 1/15/19 Central 46.00 23,000.00 22 Timothy David Home Theater 500.00 87.00 500.00 43,500.00 23 2/1/19 Central Douglas John Home Theater 24 2/18/19 Martha Alexander Home Theater 4.00 500.00 2,000.00 3/7/19 500.00 3,500.00 25 West Timothy Stephen Home Theater 7.00 3/24/19 Central Hermann 26 Video Games 50.00 58.50 2,925.00 4/10/19 Central 66.00 1,198.00 79,068.00 27 Martha Steven Television 225.00 21,600.00 28 4/27/19 Diana Cell Phone 96.00 East Martha 29 5/14/19 Central Timothy David Television 53.00 1,198.00 63,494.00 40,000.00 30 5/31/19 Central Timothy David Home Theater 80.00 500.00 6/17/19 Central Hermann 5.00 125.00 625.00 31 Shelli Desk 62.00 58.50 3,627.00 32 7/4/19 Video Games Martha Alexander Video Games 55.00 58.50 33 7/21/19 Central Hermann Sigal 3,217.50 34 8/7/19 Central Hermann Shelli Video Games 42.00 58.50 2,457.00 Stephen 3.00 125.00 35 8/24/19 West Timothy Desk 375.00 9/10/19 Central 8,386.00 36 David 7.00 1,198.00 Timothy Television Cell Phone 76.00 225.00 17,100.00 37 9/27/19 West Timothy Stephen Home Theater 57.00 500.00 28,500.00 38 10/14/19 Douglas Michael West 39 10/31/19 Central Martha Steven Television 14.00 1,198.00 16,772.00 500.00 5,500.00 40 11/17/19 Central Hermann Luis Home Theater 11.00 12/4/19 Central Hermann 94.00 500.00 Luis Home Theater 47,000.00 12/21/19 Central Steven Home Theater 28.00 42 500.00 14,000.00 Martha 1,125.00 62,550.00 pad 278.00 pad pad pad 44 pad pad pad 34.75 140.63 7,818.75 pad pad In [15]: data4=data.fillna(value='bfill') data4 OrderDate Region Manager SalesMan Units Unit_price Sale_amt 1,198.00 113,810.00 1/6/18 95.00 East Martha Alexander Television 1/23/18 Central Hermann 500.00 25,000.00 Shelli Home Theater 50.00 Television 36.00 1,198.00 43,128.00 2/9/18 Central Hermann 2/26/18 Central Timothy David Cell Phone 27.00 225.00 6,075.00 56.00 1,198.00 67,088.00 3/15/18 West Timothy Stephen Television 4/1/18 East Martha Alexander Home Theater 60.00 500.00 30,000.00 4/18/18 Central Television 75.00 1,198.00 89,850.00 Martha Steven 5/5/18 Central Hermann Luis Television 90.00 1,198.00 107,820.00 8 5/22/18 West Douglas Michael Television 32.00 1,198.00 38,336.00 500.00 30,000.00 6/8/18 East Martha Alexander Home Theater 60.00 6/25/18 Central Hermann 10 90.00 1,198.00 107,820.00 Sigal Television 11 7/12/18 East Martha Diana Home Theater 29.00 500.00 14,500.00 12 7/29/18 Karen Home Theater 81.00 500.00 40,500.00 East Douglas 13 8/15/18 East Martha Alexander Television 35.00 1,198.00 41,930.00 9/1/18 Central Douglas 2.00 125.00 250.00 14 John Desk Martha Alexander Video Games East 16.00 58.50 936.00 10/5/18 Central Hermann 16 28.00 500.00 14,000.00 Sigal Home Theater East Martha Alexander 14,400.00 **17** 10/22/18 Cell Phone 64.00 225.00 18 11/8/18 East Douglas Karen Cell Phone 15.00 225.00 3,375.00 11/25/18 Central Hermann Shelli Video Games 96.00 58.50 5,616.00 20 12/12/18 Central 67.00 1,198.00 80,266.00 Douglas John Television 12/29/18 East Douglas Karen Video Games 74.00 58.50 4,329.00 22 1/15/19 Central 46.00 500.00 23,000.00 Timothy David Home Theater 2/1/19 Central Douglas John Home Theater 87.00 500.00 43,500.00 2/18/19 4.00 500.00 2,000.00 24 East Martha Alexander Home Theater Stephen Home Theater 3,500.00 25 West Timothy 7.00 500.00 3/24/19 Central Hermann Luis Video Games 50.00 58.50 2,925.00 26 27 4/10/19 Central Martha Steven Television 66.00 1,198.00 79,068.00 28 4/27/19 East Martha Diana Cell Phone 96.00 225.00 21,600.00 29 5/14/19 Central Timothy David Television 53.00 1,198.00 63,494.00 5/31/19 Central Timothy 30 David Home Theater 80.00 500.00 40,000.00 6/17/19 Central Hermann Desk 5.00 125.00 625.00 32 7/4/19 Video Games 62.00 58.50 3,627.00 Martha Alexander East 7/21/19 Central Hermann Sigal Video Games 55.00 58.50 3,217.50 Video Games 42.00 58.50 2,457.00 34 8/7/19 Central Hermann Shelli West Timothy Stephen Desk 3.00 125.00 375.00 36 9/10/19 Central 7.00 1,198.00 8,386.00 Timothy David Television Stephen 9/27/19 West Timothy Cell Phone 76.00 225.00 17,100.00 57.00 500.00 28,500.00 10/14/19 Douglas Michael Home Theater West 10/31/19 Central Martha Steven Television 14.00 1,198.00 16,772.00 11/17/19 Central Hermann 11.00 500.00 5,500.00 40 Luis Home Theater 12/4/19 Central Hermann Luis Home Theater 94.00 500.00 47,000.00 12/21/19 Central Steven Home Theater 28.00 500.00 14,000.00 Martha bfill bfill bfill 278.00 1,125.00 62,550.00 44 bfill bfill bfill bfill bfill 34.75 140.63 7,818.75 In [16]: import numpy as np from scipy import stats In [17]: #detect outliers using IQR data2.columns Out[17]: Index(['OrderDate', 'Region', 'Manager', 'SalesMan', 'Item', 'Units', ' Unit_price ', 'Sale_amt'], dtype='object') In [20]: data2.drop(['OrderDate'], axis=1, inplace=True) data2.columns Out[20]: Index(['Region', 'Manager', 'SalesMan', 'Item', 'Units', ' Unit_price ', 'Sale_amt'], dtype='object') In [56]: data2.describe() Out[56]: Units Sale_amt 4.000000 4.000000 count 6.500000 546.500000 6.454972 302.873461 2.000000 250.000000 2.750000 343.750000 **50%** 4.000000 500.000000

