The CONTENTS Procedure

Data Set Name	BANPROJS.WALMART_FEATURES	**Observations**	8190
Member Type	DATA	**Variables**	12
Engine	V9	**Indexes**	0
Created	07/13/2021 20:34:07	**Observation Length**	144
Last Modified	07/13/2021 20:34:07	**Deleted Observations**	0
Protection		**Compressed**	NO
Data Set Type		**Sorted**	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	10
First Data Page	1
Max Obs per Page	909
Obs in First Data Page	884
Number of Data Set Repairs	0
Filename	/home/u54770142/walmart_features.sas7bdat
Release Created	9.0401M6
Host Created	Linux
Inode Number	17573044181
Access Permission	rw-r--r--
Owner Name	u54770142
File Size	1MB
File Size (bytes)	1441792

Variables in Creation Order

#	Variable	Type	Len	Format	Informat	Label	
1	Store	Num	8	BEST12.		Store	Change to Char
2	Date	Num	8	MMDDYY10.		Date	
3	Temperature	Num	8	BEST12.		Temperature	
4	Fuel_Price	Num	8	BEST12.		Fuel_Price	
5	MarkDown1	Char	14	$14.	$14.	MarkDown1	Change to Num for all Mardowns
6	MarkDown2	Char	14	$14.	$14.	MarkDown2	
7	MarkDown3	Char	14	$14.	$14.	MarkDown3	
8	MarkDown4	Char	14	$14.	$14.	MarkDown4	
9	MarkDown5	Char	14	$14.	$14.	MarkDown5	
10	CPI	Char	14	$14.	$14.	CPI	Change to Num
11	Unemployment	Char	14	$14.	$14.	Unemployment	Change to Num
12	IsHoliday	Num	8	BEST12.		IsHoliday	

Listing the first 100 Observations or WALMART_FEATURES Data Set

Obs	Store	Date	Temperature	Fuel_Price	MarkDown1	MarkDown2	MarkDown3	MarkDown4	MarkDown5	CPI	Unemployment	IsHoliday
1	1	02/05/2010	42.31	2.572	NA	NA	NA	NA	NA	211.0963582	8.106	0
2	1	02/12/2010	38.51	2.548	NA	NA	NA	NA	NA	211.2421698	8.106	1
3	1	02/19/2010	39.93	2.514	NA	NA	NA	NA	NA	211.2891429	8.106	0
4	1	02/26/2010	46.63	2.561	NA	NA	NA	NA	NA	211.3196429	8.106	0
5	1	03/05/2010	46.5	2.625	NA	NA	NA	NA	NA	211.3501429	8.106	0
6	1	03/12/2010	57.79	2.667	NA	NA	NA	NA	NA	211.3806429	8.106	0
7	1	03/19/2010	54.58	2.72	NA	NA	NA	NA	NA	211.215635	8.106	0
8	1	03/26/2010	51.45	2.732	NA	NA	NA	NA	NA	211.0180424	8.106	0
9	1	04/02/2010	62.27	2.719	NA	NA	NA	NA	NA	210.8204499	7.808	0
10	1	04/09/2010	65.86	2.77	NA	NA	NA	NA	NA	210.6228574	7.808	0
11	1	04/16/2010	66.32	2.808	NA	NA	NA	NA	NA	210.4887	7.808	0

Obs	Store	Date	Temperature	Fuel_Price	MarkDown1	MarkDown2	MarkDown3	MarkDown4	MarkDown5	CPI	Unemployment	IsHoliday
12	1	04/23/2010	64.84	2.795	NA	NA	NA	NA	NA	210.4391228	7.808	0
13	1	04/30/2010	67.41	2.78	NA	NA	NA	NA	NA	210.3895456	7.808	0
14	1	05/07/2010	72.55	2.835	NA	NA	NA	NA	NA	210.3399684	7.808	0
15	1	05/14/2010	74.78	2.854	NA	NA	NA	NA	NA	210.3374261	7.808	0
16	1	05/21/2010	76.44	2.826	NA	NA	NA	NA	NA	210.6170934	7.808	0
17	1	05/28/2010	80.44	2.759	NA	NA	NA	NA	NA	210.8967606	7.808	0
18	1	06/04/2010	80.69	2.705	NA	NA	NA	NA	NA	211.1764278	7.808	0
19	1	06/11/2010	80.43	2.668	NA	NA	NA	NA	NA	211.4560951	7.808	0
20	1	06/18/2010	84.11	2.637	NA	NA	NA	NA	NA	211.4537719	7.808	0
21	1	06/25/2010	84.34	2.653	NA	NA	NA	NA	NA	211.3386526	7.808	0
22	1	07/02/2010	80.91	2.669	NA	NA	NA	NA	NA	211.2235333	7.787	0
23	1	07/09/2010	80.48	2.642	NA	NA	NA	NA	NA	211.108414	7.787	0
24	1	07/16/2010	83.15	2.623	NA	NA	NA	NA	NA	211.1003854	7.787	0
25	1	07/23/2010	83.36	2.608	NA	NA	NA	NA	NA	211.2351443	7.787	0
26	1	07/30/2010	81.84	2.64	NA	NA	NA	NA	NA	211.3699032	7.787	0
27	1	08/06/2010	87.16	2.627	NA	NA	NA	NA	NA	211.5046621	7.787	0
28	1	08/13/2010	87	2.692	NA	NA	NA	NA	NA	211.6394211	7.787	0
29	1	08/20/2010	86.65	2.664	NA	NA	NA	NA	NA	211.6033633	7.787	0
30	1	08/27/2010	85.22	2.619	NA	NA	NA	NA	NA	211.5673056	7.787	0
31	1	09/03/2010	81.21	2.577	NA	NA	NA	NA	NA	211.5312479	7.787	0
32	1	09/10/2010	78.69	2.565	NA	NA	NA	NA	NA	211.4951902	7.787	1
33	1	09/17/2010	82.11	2.582	NA	NA	NA	NA	NA	211.5224596	7.787	0
34	1	09/24/2010	80.94	2.624	NA	NA	NA	NA	NA	211.5972246	7.787	0
35	1	10/01/2010	71.89	2.603	NA	NA	NA	NA	NA	211.6719895	7.838	0
36	1	10/08/2010	63.93	2.633	NA	NA	NA	NA	NA	211.7467544	7.838	0
37	1	10/15/2010	67.18	2.72	NA	NA	NA	NA	NA	211.8137436	7.838	0
38	1	10/22/2010	69.86	2.725	NA	NA	NA	NA	NA	211.8612937	7.838	0
39	1	10/29/2010	69.64	2.716	NA	NA	NA	NA	NA	211.9088438	7.838	0
40	1	11/05/2010	58.74	2.689	NA	NA	NA	NA	NA	211.9563939	7.838	0
41	1	11/12/2010	59.61	2.728	NA	NA	NA	NA	NA	212.003944	7.838	0
42	1	11/19/2010	51.41	2.771	NA	NA	NA	NA	NA	211.8896737	7.838	0
43	1	11/26/2010	64.52	2.735	NA	NA	NA	NA	NA	211.7484333	7.838	1
44	1	12/03/2010	49.27	2.708	NA	NA	NA	NA	NA	211.607193	7.838	0
45	1	12/10/2010	46.33	2.843	NA	NA	NA	NA	NA	211.4659526	7.838	0
46	1	12/17/2010	49.84	2.869	NA	NA	NA	NA	NA	211.4053124	7.838	0
47	1	12/24/2010	52.33	2.886	NA	NA	NA	NA	NA	211.4051222	7.838	0
48	1	12/31/2010	48.43	2.943	NA	NA	NA	NA	NA	211.4049321	7.838	1
49	1	01/07/2011	48.27	2.976	NA	NA	NA	NA	NA	211.4047419	7.742	0
50	1	01/14/2011	35.4	2.983	NA	NA	NA	NA	NA	211.4574109	7.742	0
51	1	01/21/2011	44.04	3.016	NA	NA	NA	NA	NA	211.8272343	7.742	0
52	1	01/28/2011	43.83	3.01	NA	NA	NA	NA	NA	212.1970577	7.742	0
53	1	02/04/2011	42.27	2.989	NA	NA	NA	NA	NA	212.5668812	7.742	0
54	1	02/11/2011	36.39	3.022	NA	NA	NA	NA	NA	212.9367046	7.742	1
55	1	02/18/2011	57.36	3.045	NA	NA	NA	NA	NA	213.2478853	7.742	0
56	1	02/25/2011	62.9	3.065	NA	NA	NA	NA	NA	213.535609	7.742	0
57	1	03/04/2011	59.58	3.288	NA	NA	NA	NA	NA	213.8233327	7.742	0
58	1	03/11/2011	53.56	3.459	NA	NA	NA	NA	NA	214.1110564	7.742	0
59	1	03/18/2011	62.76	3.488	NA	NA	NA	NA	NA	214.3627114	7.742	0
60	1	03/25/2011	69.97	3.473	NA	NA	NA	NA	NA	214.5999389	7.742	0
61	1	04/01/2011	59.17	3.524	NA	NA	NA	NA	NA	214.8371664	7.682	0
62	1	04/08/2011	67.84	3.622	NA	NA	NA	NA	NA	215.0743939	7.682	0
63	1	04/15/2011	71.27	3.743	NA	NA	NA	NA	NA	215.2918561	7.682	0
64	1	04/22/2011	72.99	3.807	NA	NA	NA	NA	NA	215.4599053	7.682	0
65	1	04/29/2011	72.03	3.81	NA	NA	NA	NA	NA	215.6279544	7.682	0
66	1	05/06/2011	64.61	3.906	NA	NA	NA	NA	NA	215.7960035	7.682	0
67	1	05/13/2011	75.64	3.899	NA	NA	NA	NA	NA	215.9640526	7.682	0
68	1	05/20/2011	67.63	3.907	NA	NA	NA	NA	NA	215.7339202	7.682	0
69	1	05/27/2011	77.72	3.786	NA	NA	NA	NA	NA	215.5037878	7.682	0

Obs	Store	Date	Temperature	Fuel_Price	MarkDown1	MarkDown2	MarkDown3	MarkDown4	MarkDown5	CPI	Unemployment	IsHoliday
70	1	06/03/2011	83	3.699	NA	NA	NA	NA	NA	215.2736553	7.682	0
71	1	06/10/2011	83.13	3.648	NA	NA	NA	NA	NA	215.0435229	7.682	0
72	1	06/17/2011	86.41	3.637	NA	NA	NA	NA	NA	214.9980596	7.682	0
73	1	06/24/2011	83.58	3.594	NA	NA	NA	NA	NA	215.0910982	7.682	0
74	1	07/01/2011	85.55	3.524	NA	NA	NA	NA	NA	215.1841368	7.962	0
75	1	07/08/2011	85.83	3.48	NA	NA	NA	NA	NA	215.2771754	7.962	0
76	1	07/15/2011	88.54	3.575	NA	NA	NA	NA	NA	215.3611087	7.962	0
77	1	07/22/2011	85.77	3.651	NA	NA	NA	NA	NA	215.4222784	7.962	0
78	1	07/29/2011	86.83	3.682	NA	NA	NA	NA	NA	215.4834482	7.962	0
79	1	08/05/2011	91.65	3.684	NA	NA	NA	NA	NA	215.544618	7.962	0
80	1	08/12/2011	90.76	3.638	NA	NA	NA	NA	NA	215.6057878	7.962	0
81	1	08/19/2011	89.94	3.554	NA	NA	NA	NA	NA	215.6693107	7.962	0
82	1	08/26/2011	87.96	3.523	NA	NA	NA	NA	NA	215.7332258	7.962	0
83	1	09/02/2011	87.83	3.533	NA	NA	NA	NA	NA	215.7971409	7.962	0
84	1	09/09/2011	76	3.546	NA	NA	NA	NA	NA	215.861056	7.962	1
85	1	09/16/2011	79.94	3.526	NA	NA	NA	NA	NA	216.0410526	7.962	0
86	1	09/23/2011	75.8	3.467	NA	NA	NA	NA	NA	216.3758246	7.962	0
87	1	09/30/2011	79.69	3.355	NA	NA	NA	NA	NA	216.7105965	7.962	0
88	1	10/07/2011	69.31	3.285	NA	NA	NA	NA	NA	217.0453684	7.866	0
89	1	10/14/2011	71.74	3.274	NA	NA	NA	NA	NA	217.3552733	7.866	0
90	1	10/21/2011	63.71	3.353	NA	NA	NA	NA	NA	217.5159762	7.866	0
91	1	10/28/2011	66.57	3.372	NA	NA	NA	NA	NA	217.6766791	7.866	0
92	1	11/04/2011	54.98	3.332	NA	NA	NA	NA	NA	217.837382	7.866	0
93	1	11/11/2011	59.11	3.297	10382.9	6115.67	215.07	2406.62	6551.42	217.9980849	7.866	0
94	1	11/18/2011	62.25	3.308	6074.12	254.39	51.98	427.39	5988.57	218.2205088	7.866	0
95	1	11/25/2011	60.14	3.236	410.31	98	55805.51	8	554.92	218.4676211	7.866	1
96	1	12/02/2011	48.91	3.172	5629.51	68	1398.11	2084.64	20475.32	218.7147333	7.866	0
97	1	12/09/2011	43.93	3.158	4640.65	19	105.02	3639.42	14461.82	218.9618456	7.866	0
98	1	12/16/2011	51.63	3.159	5011.32	67	347.37	225.79	4011.37	219.1794533	7.866	0
99	1	12/23/2011	47.96	3.112	2725.36	40.48	634.7	24.9	2739.43	219.3577216	7.866	0
100	1	12/30/2011	44.55	3.129	5762.1	46011.38	260.36	983.65	4735.78	219.5359898	7.866	1

Checking Missing Values for Markdown1-5 Variables

The FREQ Procedure

MarkDown1		
MarkDown1	Frequency	Percent
Nonmissing	4032	49.23
Missing	4158	50.77

Analyze values after Nov 2011 only - do we remove the other rows completely? or ue them to analyza impact of CPI/ Temperature etc on Sales? My opinion is to leave those rows to analyze against other dependant variables

Create Binary Variable IsMarkdown for all Mardown1-5 together, Leave markdown1-5 in datset but change it to num.

MarkDown2		
MarkDown2	Frequency	Percent
Nonmissing	2921	35.67
Missing	5269	64.33

Create variable : Number of Categories on

Markdown Example: 5 (categories on markdown:5, 2

(2 categories on markdown)

MarkDown3		
MarkDown3	Frequency	Percent
Nonmissing	3613	44.11
Missing	4577	55.89

MarkDown4		
MarkDown4	Frequency	Percent
Nonmissing	3464	42.30
Missing	4726	57.70

MarkDown5		
MarkDown5	Frequency	Percent
Nonmissing	4050	49.45
Missing	4140	50.55

Checking Missing Values for CPI

The FREQ Procedure

CPI		
CPI	Frequency	Percent
Nonmissing	7605	92.86
Missing	585	7.14

Create Ordinal Variable - High/Medium/Low

Checking Missing Values for Unemployment

The FREQ Procedure

Unemployment		
Unemployment	Frequency	Percent
Nonmissing	7605	92.86
Missing	585	7.14

Checking Numeric Missing Values for Numeric Variables in WALMART_FEATURES dataset

The MEANS Procedure

Variable	Label	N	N Miss	Mean	Median	Minimum	Maximum
Store	Store	8190	0	23.0000000	23.0000000	1.0000000	45.0000000
Date	Date	8190	0	18931.50	18931.50	18298.00	19565.00
Temperature	Temperature	8190	0	59.3561978	60.7100000	-7.2900000	101.9500000
Fuel_Price	Fuel_Price	8190	0	3.4059918	3.5130000	2.4720000	4.4680000
IsHoliday	IsHoliday	8190	0	0.0714286	0	0	1.0000000

Create Temperature Zone (categorize store): NorthWest SouthWest Midwest.....

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Variable: Store (Store)

Moments			
N	8190	Sum Weights	8190
Mean	23	Sum Observations	188370
Std Deviation	12.9879661	Variance	168.687263
Skewness	0	Kurtosis	-1.2011865
Uncorrected SS	5713890	Corrected SS	1381380
Coeff Variation	56.4694178	Std Error Mean	0.14351563

Basic Statistical Measures			
Location		Variability	
Mean	23.00000	Std Deviation	12.98797
Median	23.00000	Variance	168.68726
Mode	1.00000	Range	44.00000
		Interquartile Range	22.00000

Note: The mode displayed is the smallest of 45 modes with a count of 182.

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	160.2613	Pr > |t|	<.0001
Sign	M	4095	Pr >= |M|	<.0001

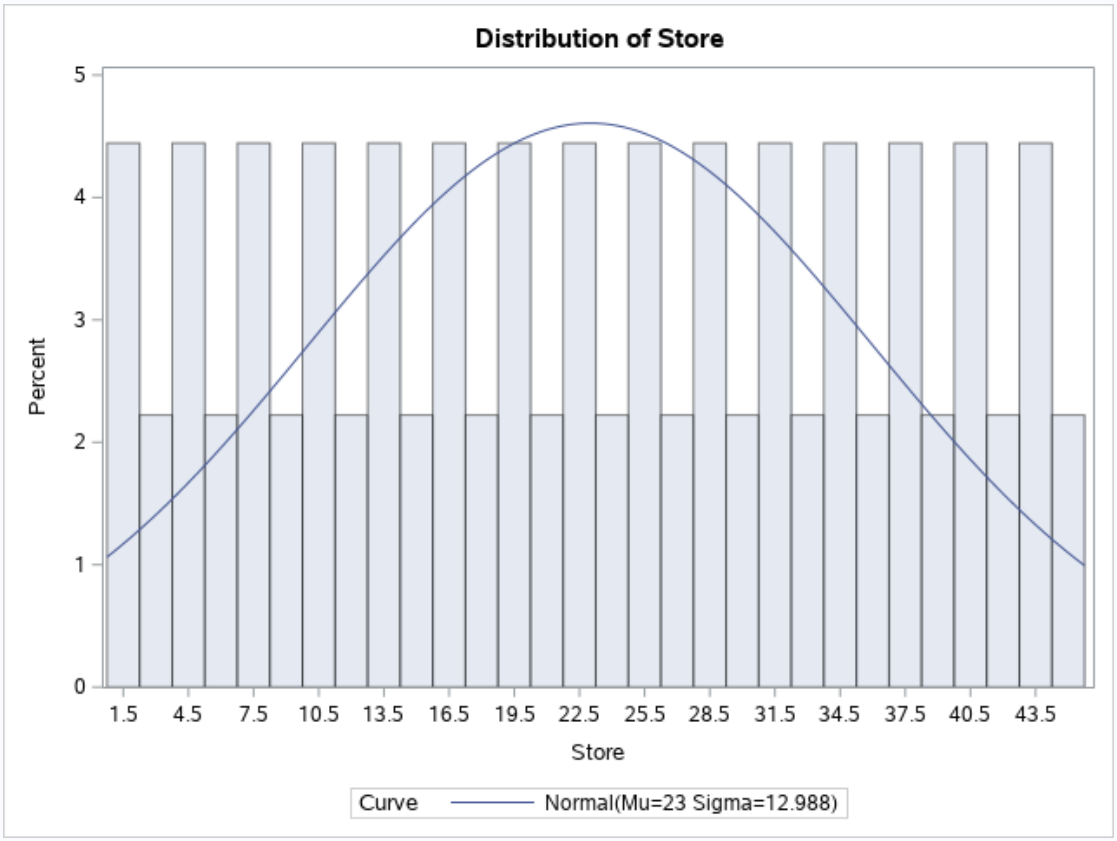
Tests for Location: Mu0=0				
Test	Statistic		p Value	
Signed Rank	S	16771073	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	45
99%	45
95%	43
90%	41
75% Q3	34
50% Median	23
25% Q1	12
10%	5
5%	3
1%	1
0% Min	1

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	182	45	8186
1	181	45	8187
1	180	45	8188
1	179	45	8189
1	178	45	8190

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure



Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Fitted Normal Distribution for Store (Store)

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	23
Std Dev	Sigma	12.98797

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.0682222	Pr > D	<0.010
Cramer-von Mises	W-Sq	12.8257705	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	93.5057437	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
1.0	1.00000	-7.21453
5.0	3.00000	1.63670
10.0	5.00000	6.35525
25.0	12.00000	14.23975
50.0	23.00000	23.00000
75.0	34.00000	31.76025
90.0	41.00000	39.64475
95.0	43.00000	44.36330
99.0	45.00000	53.21453

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Variable: Date (Date)

Moments			
N	8190	Sum Weights	8190
Mean	18931.5	Sum Observations	155048985
Std Deviation	367.789024	Variance	135268.766
Skewness	0	Kurtosis	-1.2000725
Uncorrected SS	2.93642E12	Corrected SS	1107715928
Coeff Variation	1.94273578	Std Error Mean	4.06402914

Basic Statistical Measures			
Location		Variability	
Mean	18931.50	Std Deviation	367.78902
Median	18931.50	Variance	135269
Mode	18298.00	Range	1267
		Interquartile Range	637.00000

Note: The mode displayed is the smallest of 182 modes with a count of 45.

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	4658.308	Pr > |t|	<.0001
Sign	M	4095	Pr >= |M|	<.0001
Signed Rank	S	16771073	Pr >= |S|	<.0001

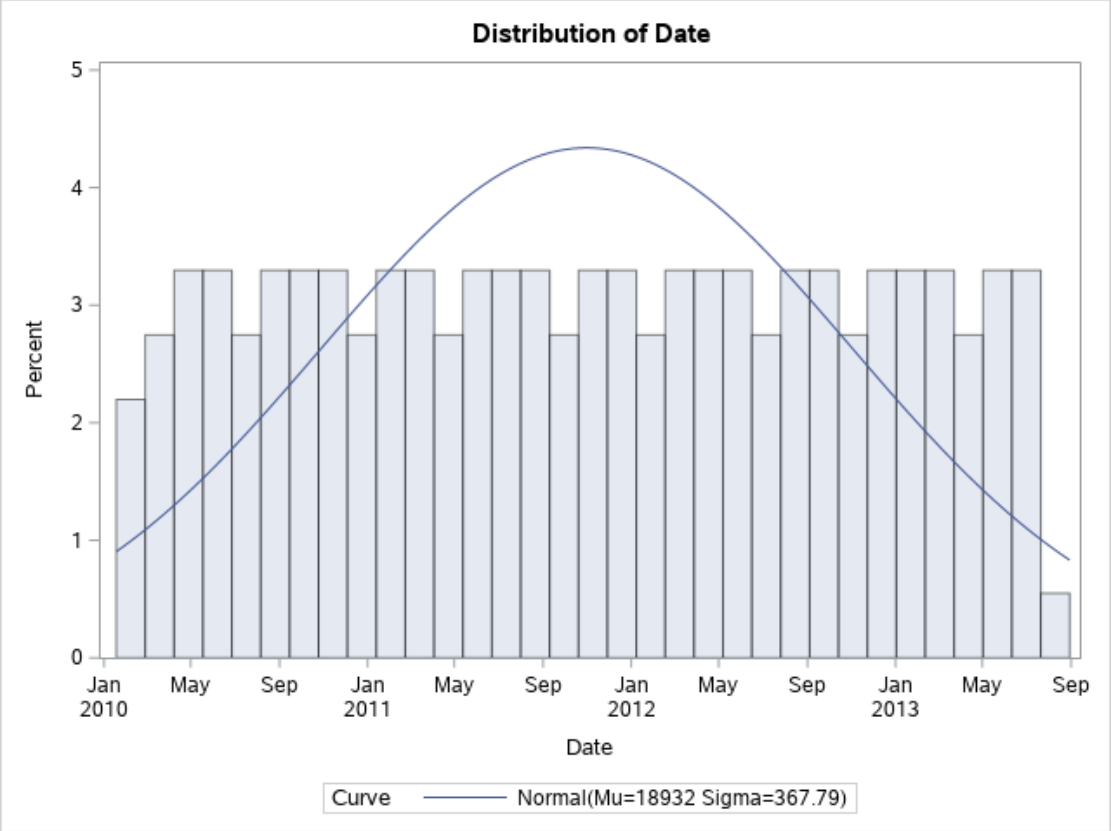
Quantiles (Definition 5)	
Level	Quantile
100% Max	19565.0
99%	19558.0
95%	19502.0

Quantiles (Definition 5)	
Level	Quantile
90%	19439.0
75% Q3	19250.0
50% Median	18931.5
25% Q1	18613.0
10%	18424.0
5%	18361.0
1%	18305.0
0% Min	18298.0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
18298	8009	19565	7462
18298	7827	19565	7644
18298	7645	19565	7826
18298	7463	19565	8008
18298	7281	19565	8190

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure



Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Fitted Normal Distribution for Date (Date)

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	18931.5
Std Dev	Sigma	367.789

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.0599409	Pr > D	<0.010
Cramer-von Mises	W-Sq	12.5046648	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	91.1838632	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
1.0	18305.0	18075.9
5.0	18361.0	18326.5
10.0	18424.0	18460.2
25.0	18613.0	18683.4
50.0	18931.5	18931.5
75.0	19250.0	19179.6
90.0	19439.0	19402.8
95.0	19502.0	19536.5
99.0	19558.0	19787.1

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Variable: Temperature (Temperature)

Moments			
N	8190	Sum Weights	8190
Mean	59.3561978	Sum Observations	486127.26
Std Deviation	18.6786068	Variance	348.890354
Skewness	-0.2833844	Kurtosis	-0.6108838
Uncorrected SS	31711728.9	Corrected SS	2857063.11
Coeff Variation	31.4686714	Std Error Mean	0.2063966

Basic Statistical Measures			
Location		Variability	
Mean	59.35620	Std Deviation	18.67861
Median	60.71000	Variance	348.89035
Mode	50.43000	Range	109.24000
		Interquartile Range	27.98000

Note: The mode displayed is the smallest of 2 modes with a count of 11.

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	287.5832	Pr > |t|	<.0001
Sign	M	4091	Pr >= |M|	<.0001
Signed Rank	S	16771043	Pr >= |S|	<.0001

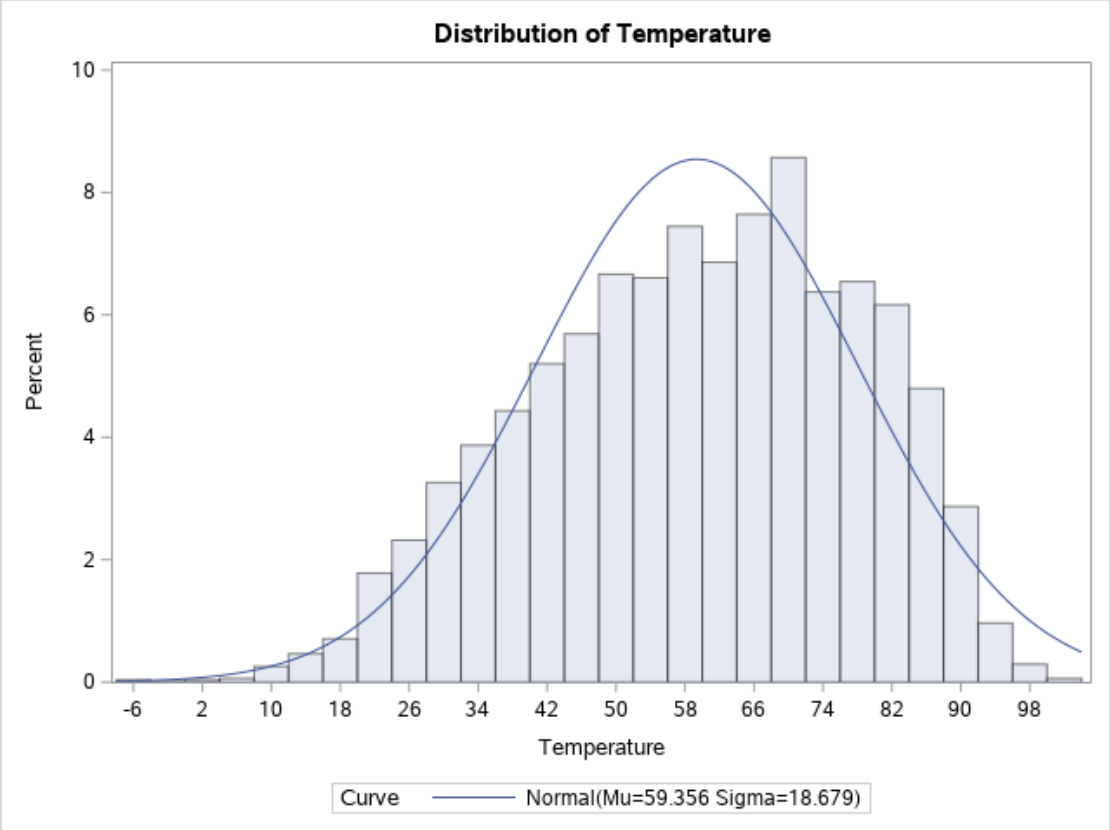
Quantiles (Definition 5)	
Level	Quantile
100% Max	101.95
99%	92.98
95%	87.14
90%	83.36
75% Q3	73.88
50% Median	60.71
25% Q1	45.90
10%	33.25
5%	26.84
1%	17.00

Quantiles (Definition 5)	
Level	Quantile
0% Min	-7.29

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
-7.29	1247	100.07	5956
-6.61	1246	100.14	5848
-6.08	1245	101.95	2181
-2.06	1145	101.95	5093
0.25	3067	101.95	6913

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure



Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Fitted Normal Distribution for Temperature (Temperature)

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	59.3562
Std Dev	Sigma	18.67861

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.0464662	Pr > D	<0.010
Cramer-von Mises	W-Sq	4.9922664	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	33.2941460	Pr > A-Sq	<0.005

Quantiles for Normal Distribution

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	17.0000	15.9033
5.0	26.8400	28.6326
10.0	33.2500	35.4186
25.0	45.9000	46.7577
50.0	60.7100	59.3562
75.0	73.8800	71.9547
90.0	83.3600	83.2938
95.0	87.1400	90.0798
99.0	92.9800	102.8091

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Variable: Fuel_Price (Fuel_Price)

Moments			
N	8190	Sum Weights	8190
Mean	3.40599182	Sum Observations	27895.073
Std Deviation	0.43133657	Variance	0.18605124
Skewness	-0.3050626	Kurtosis	-0.9523877
Uncorrected SS	96533.964	Corrected SS	1523.57358
Coeff Variation	12.6640519	Std Error Mean	0.00476622

DEcide whether we create a ordincal variables: High, medium, low

Basic Statistical Measures			
Location		Variability	
Mean	3.405992	Std Deviation	0.43134
Median	3.513000	Variance	0.18605
Mode	3.417000	Range	1.99600
		Interquartile Range	0.70200

Note: The mode displayed is the smallest of 2 modes with a count of 43.

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	714.6103	Pr > |t|	<.0001
Sign	M	4095	Pr >= |M|	<.0001
Signed Rank	S	16771073	Pr >= |S|	<.0001

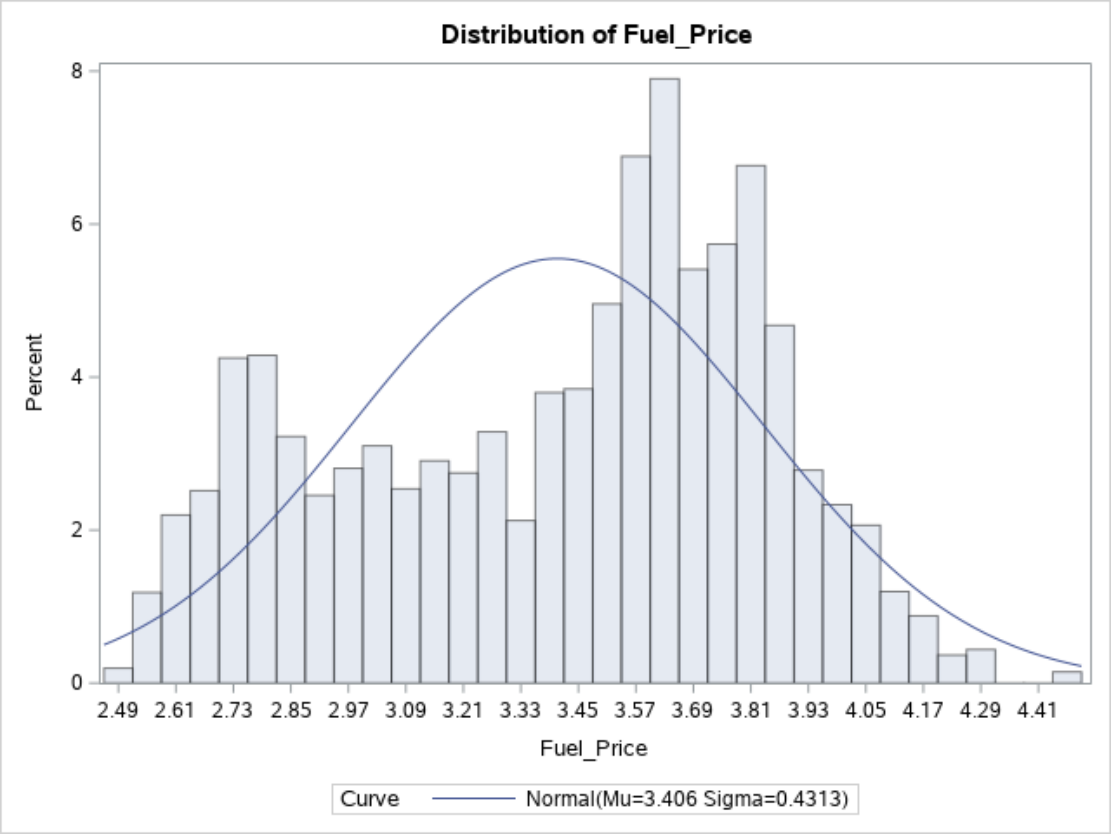
Quantiles (Definition 5)	
Level	Quantile
100% Max	4.468
99%	4.193
95%	4.021
90%	3.903
75% Q3	3.743
50% Median	3.513
25% Q1	3.041
10%	2.755
5%	2.669
1%	2.572
0% Min	2.472

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
2.472	6373	4.468	2143
2.513	6402	4.468	5055
2.514	7647	4.468	5965
2.514	6919	4.468	6875
2.514	6555	4.468	7603

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure



Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Fitted Normal Distribution for Fuel_Price (Fuel_Price)

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	3.405992
Std Dev	Sigma	0.431337

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.103733	Pr > D	<0.010
Cramer-von Mises	W-Sq	22.612192	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	130.905646	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
1.0	2.57200	2.40255
5.0	2.66900	2.69651

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
10.0	2.75500	2.85321
25.0	3.04100	3.11506
50.0	3.51300	3.40599
75.0	3.74300	3.69692
90.0	3.90300	3.95877
95.0	4.02100	4.11548
99.0	4.19300	4.40943

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Variable: IsHoliday (IsHoliday)

Moments			
N	8190	Sum Weights	8190
Mean	0.07142857	Sum Observations	585
Std Deviation	0.2575551	Variance	0.06633463
Skewness	3.32881088	Kurtosis	9.08319994
Uncorrected SS	585	Corrected SS	543.214286
Coeff Variation	360.577141	Std Error Mean	0.00284596

Basic Statistical Measures			
Location		Variability	
Mean	0.071429	Std Deviation	0.25756
Median	0.000000	Variance	0.06633
Mode	0.000000	Range	1.00000
		Interquartile Range	0

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	25.09827	Pr > |t|	<.0001
Sign	M	292.5	Pr >= |M|	<.0001
Signed Rank	S	85702.5	Pr >= |S|	<.0001

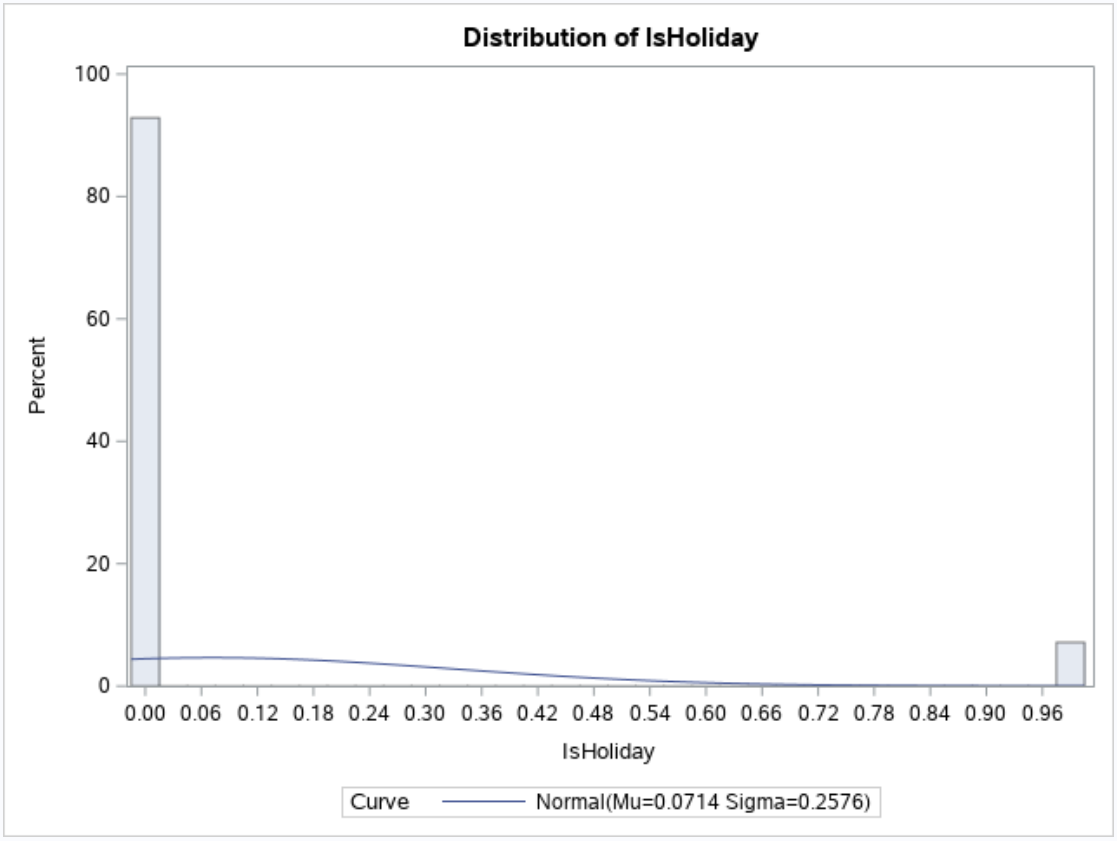
Quantiles (Definition 5)	
Level	Quantile
100% Max	1
99%	1
95%	1
90%	0
75% Q3	0
50% Median	0
25% Q1	0
10%	0
5%	0
1%	0
0% Min	0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	8190	1	8114
0	8189	1	8144
0	8188	1	8155
0	8187	1	8160

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	8186	1	8166

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure



Create a HolidayName variable, categorical

Using PROC UNIVARIATE to Examine store date temperature Fuel_price IsHoliday

The UNIVARIATE Procedure
Fitted Normal Distribution for IsHoliday (IsHoliday)

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	0.071429
Std Dev	Sigma	0.257555

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.53781	Pr > D	<0.010
Cramer-von Mises	W-Sq	588.54656	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	2849.95798	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
1.0	0.00000	-0.52773
5.0	0.00000	-0.35221
10.0	0.00000	-0.25864
25.0	0.00000	-0.10229
50.0	0.00000	0.07143
75.0	0.00000	0.24515

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
90.0	0.00000	0.40150
95.0	1.00000	0.49507
99.0	1.00000	0.67059