13 May 2024

SAS Project Report

Prepared by: Metin Mirza VATANSEVER

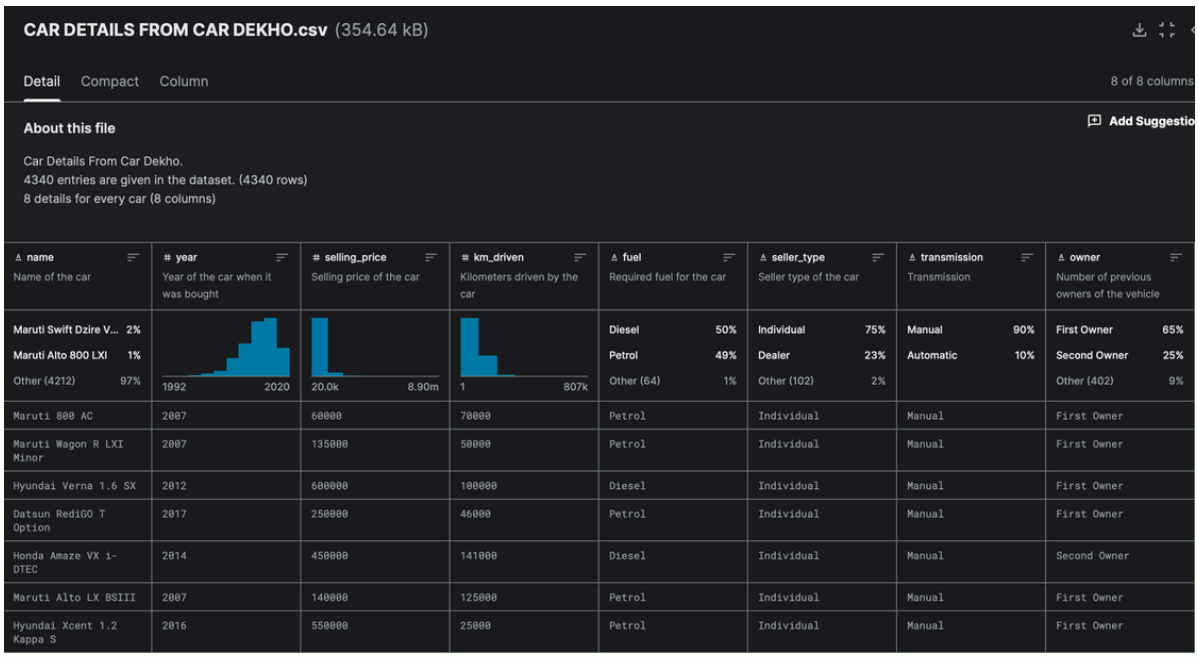
Lecture: SAS Programming
Lecture's Teacher: Gülhayat GÖLBAŞI ŞİMŞEK

About Dataset

This dataset comprises data collected from 'Car Dekho'. It includes the following details of cars:

1) Car name: The brand and model of the car.
2) Year: The manufacturing year of the car.
3) Selling Price: The price at which the car is being sold.
4) Kms driven: The distance covered by the car in kilometers.
5) Fuel: The type of fuel used by the car (petrol, diesel, CNG, etc.).
6) Seller type: The type of seller (individual or dealer).
7) Transmission: The transmission type of the car (manual or automatic).
8) Owner: The number of previous owners of the car.
9) Horsepower: Horsepower of the car

These explanations clarify what each variable represents in the dataset and what type of information it contains. This information will assist the readers of the report in understanding the dataset.



CAR DETAILS FROM CAR DEKHO.csv (354.64 kB)

Detail | Compact | Column — 8 of 8 columns

About this file

Car Details From Car Dekho.
4340 entries are given in the dataset. (4340 rows)
8 details for every car (8 columns)

name Name of the car	year Year of the car when it was bought	selling_price Selling price of the car	km_driven Kilometers driven by the car	fuel Required fuel for the car	seller_type Seller type of the car	transmission Transmission	owner Number of previous owners of the vehicle
Maruti Swift Dzire V... 2% Maruti Alto 800 LXI 1% Other (4212) 97%	1992 2020	20.0k 8.90m	1 807k	Diesel 50% Petrol 49% Other (64) 1%	Individual 75% Dealer 23% Other (102) 2%	Manual 90% Automatic 10%	First Owner 65% Second Owner 25% Other (402) 9%
Maruti 800 AC	2007	60000	70000	Petrol	Individual	Manual	First Owner
Maruti Wagon R LXI Minor	2007	135000	50000	Petrol	Individual	Manual	First Owner
Hyundai Verna 1.6 SX	2012	600000	100000	Diesel	Individual	Manual	First Owner
Datsun RediGO T Option	2017	250000	46000	Petrol	Individual	Manual	First Owner
Honda Amaze VX i-DTEC	2014	450000	141000	Diesel	Individual	Manual	Second Owner
Maruti Alto LX BSIII	2007	140000	125000	Petrol	Individual	Manual	First Owner
Hyundai Xcent 1.2 Kappa S	2016	550000	25000	Petrol	Individual	Manual	First Owner

Here is some information about the dataset

03

Code-Output

```
proc import datafile="/home/u63788634/CAR DETAILS FROM CAR
DEKHO.csv"
  out=car_data dbms=csv replace;
run;
proc datasets library=work;
 modify car_data;
 label name='Car Name' year='Year of Manufacture' selling_price='Selling
Price'
  km_driven='Kilometers Driven' fuel='Fuel Type' seller_type='Seller Type'
  transmission='Transmission Type' owner='Number of Owners';
quit;
/* Veri setini yazdırma */
```

Directory	
Libref	WORK
Engine	V9
Physical Name	/saswork/SAS_workFC380000AA6D_odaws02-euw1.oda.sas.com/SAS_work63050000AA6D_odaws02-euw1.oda.sas.com
Filename	/saswork/SAS_workFC380000AA6D_odaws02-euw1.oda.sas.com/SAS_work63050000AA6D_odaws02-euw1.oda.sas.com
Inode Number	1610701510
Access Permission	rwx------
Owner Name	u63788634
File Size	4KB
File Size (bytes)	4096

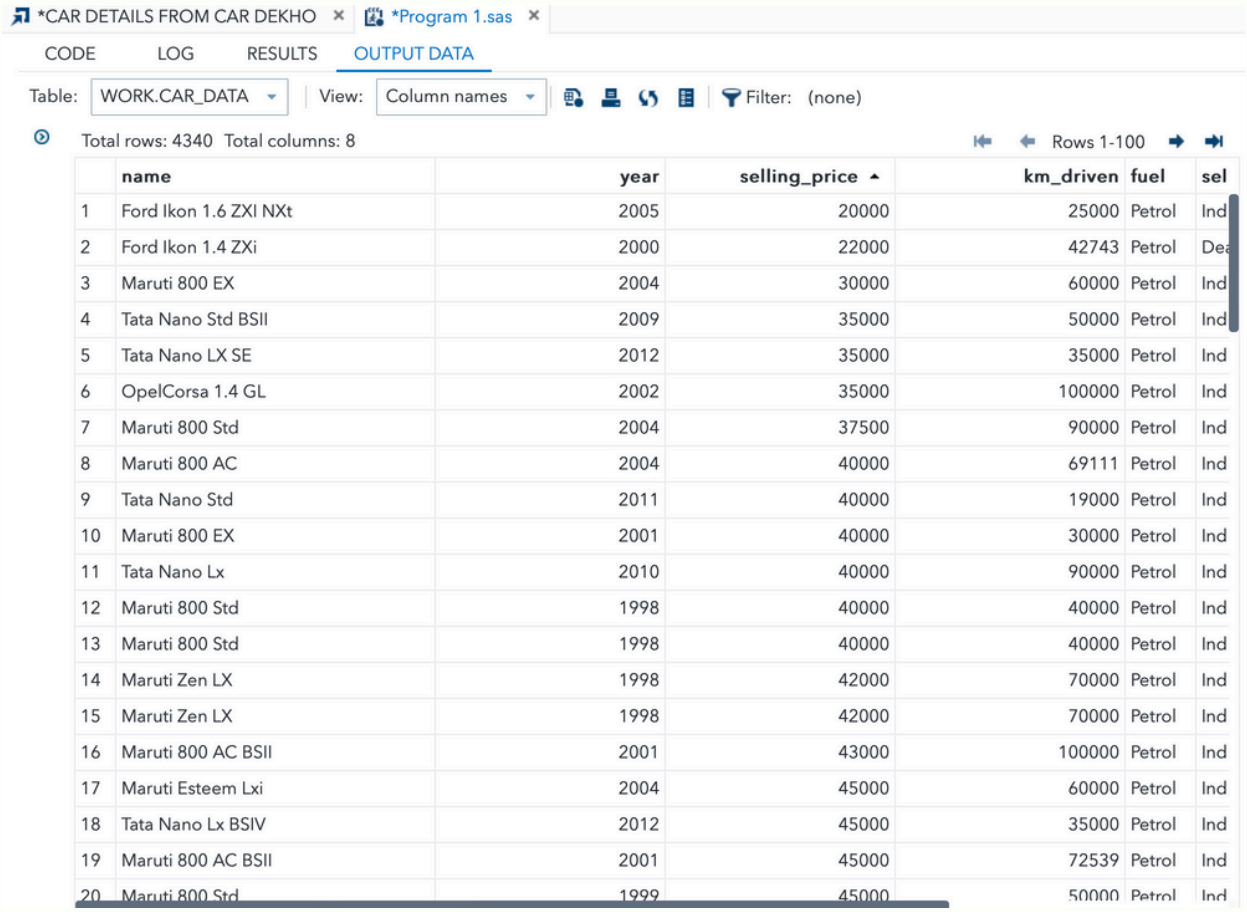
#	Name	Member Type	File Size	Last Modified
1	REGSTRY	ITEMSTOR	32KB	12/05/2024 19:19:28
2	SASGOPT	CATALOG	12KB	12/05/2024 19:19:28
3	SASMAC1	CATALOG	208KB	12/05/2024 19:19:28
4	SASMAC2	CATALOG	20KB	12/05/2024 19:19:28
5	SASMAC3	CATALOG	20KB	12/05/2024 19:19:28
6	SASMAC4	CATALOG	20KB	12/05/2024 19:20:13
7	SASMAC5	CATALOG	20KB	12/05/2024 19:19:28
8	SASMAC6	CATALOG	20KB	12/05/2024 19:19:28
9	SASMAC7	CATALOG	20KB	12/05/2024 19:19:28
10	SASMAC8	CATALOG	20KB	12/05/2024 19:19:28
11	SASMAC9	CATALOG	20KB	12/05/2024 19:19:28
12	SASMACR	CATALOG	20KB	12/05/2024 19:19:28

04

Code-Output

```
proc print data=car_data;
 title 'Car data with Labels';
run;

proc print data=car_data;
 title 'Car Dataset';
run;
proc sort data=car_data;
 by descending selling_price;
run;
```



*CAR DETAILS FROM CAR DEKHO × *Program 1.sas ×

CODE LOG RESULTS OUTPUT DATA

Table: WORK.CAR_DATA View: Column names Filter: (none)

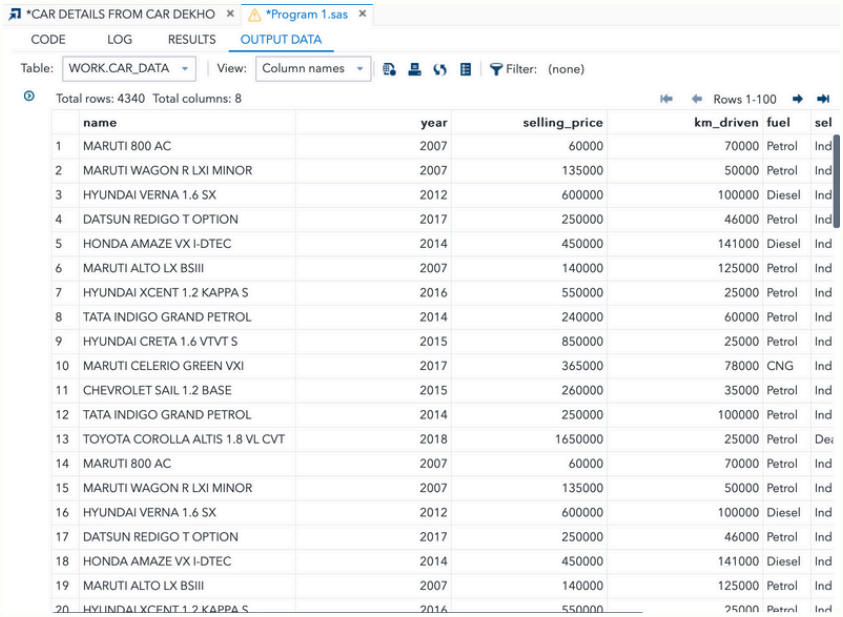
Total rows: 4340 Total columns: 8 Rows 1-100

	name	year	selling_price	km_driven	fuel	sel
1	Ford Ikon 1.6 ZXI NXt	2005	20000	25000	Petrol	Ind
2	Ford Ikon 1.4 ZXi	2000	22000	42743	Petrol	Dea
3	Maruti 800 EX	2004	30000	60000	Petrol	Ind
4	Tata Nano Std BSII	2009	35000	50000	Petrol	Ind
5	Tata Nano LX SE	2012	35000	35000	Petrol	Ind
6	OpelCorsa 1.4 GL	2002	35000	100000	Petrol	Ind
7	Maruti 800 Std	2004	37500	90000	Petrol	Ind
8	Maruti 800 AC	2004	40000	69111	Petrol	Ind
9	Tata Nano Std	2011	40000	19000	Petrol	Ind
10	Maruti 800 EX	2001	40000	30000	Petrol	Ind
11	Tata Nano Lx	2010	40000	90000	Petrol	Ind
12	Maruti 800 Std	1998	40000	40000	Petrol	Ind
13	Maruti 800 Std	1998	40000	40000	Petrol	Ind
14	Maruti Zen LX	1998	42000	70000	Petrol	Ind
15	Maruti Zen LX	1998	42000	70000	Petrol	Ind
16	Maruti 800 AC BSII	2001	43000	100000	Petrol	Ind
17	Maruti Esteem Lxi	2004	45000	60000	Petrol	Ind
18	Tata Nano Lx BSIV	2012	45000	35000	Petrol	Ind
19	Maruti 800 AC BSII	2001	45000	72539	Petrol	Ind
20	Maruti 800 Std	1999	45000	50000	Petrol	Ind

05

Code-Output

```
proc print data=car_data;
 format selling_price dollar12.2;
 title 'Car Dataset Using DOLLAR Format';
run;
data car_data;
 set car_data;
 name=upcase(name);
 title 'Car Data with Upcase Function' run;
```



*CAR DETAILS FROM CAR DEKHO | *Program 1.sas

CODE | LOG | RESULTS | OUTPUT DATA

Table: WORK.CAR_DATA | View: Column names | Filter: (none)

Total rows: 4340 Total columns: 8 | Rows 1-100

	name	year	selling_price	km_driven	fuel	sel
1	MARUTI 800 AC	2007	60000	70000	Petrol	Ind
2	MARUTI WAGON R LXI MINOR	2007	135000	50000	Petrol	Ind
3	HYUNDAI VERNA 1.6 SX	2012	600000	100000	Diesel	Ind
4	DATSUN REDIGO T OPTION	2017	250000	46000	Petrol	Ind
5	HONDA AMAZE VX I-DTEC	2014	450000	141000	Diesel	Ind
6	MARUTI ALTO LX BSIII	2007	140000	125000	Petrol	Ind
7	HYUNDAI XCENT 1.2 KAPPA S	2016	550000	25000	Petrol	Ind
8	TATA INDIGO GRAND PETROL	2014	240000	60000	Petrol	Ind
9	HYUNDAI CRETA 1.6 VTVT S	2015	850000	25000	Petrol	Ind
10	MARUTI CELERIO GREEN VXI	2017	365000	78000	CNG	Ind
11	CHEVROLET SAIL 1.2 BASE	2015	260000	35000	Petrol	Ind
12	TATA INDIGO GRAND PETROL	2014	250000	100000	Petrol	Ind
13	TOYOTA COROLLA ALTIS 1.8 VL CVT	2018	1650000	25000	Petrol	Dea
14	MARUTI 800 AC	2007	60000	70000	Petrol	Ind
15	MARUTI WAGON R LXI MINOR	2007	135000	50000	Petrol	Ind
16	HYUNDAI VERNA 1.6 SX	2012	600000	100000	Diesel	Ind
17	DATSUN REDIGO T OPTION	2017	250000	46000	Petrol	Ind
18	HONDA AMAZE VX I-DTEC	2014	450000	141000	Diesel	Ind
19	MARUTI ALTO LX BSIII	2007	140000	125000	Petrol	Ind
20	HYUNDAI XCENT 1.2 KAPPA S	2016	550000	25000	Petrol	Ind

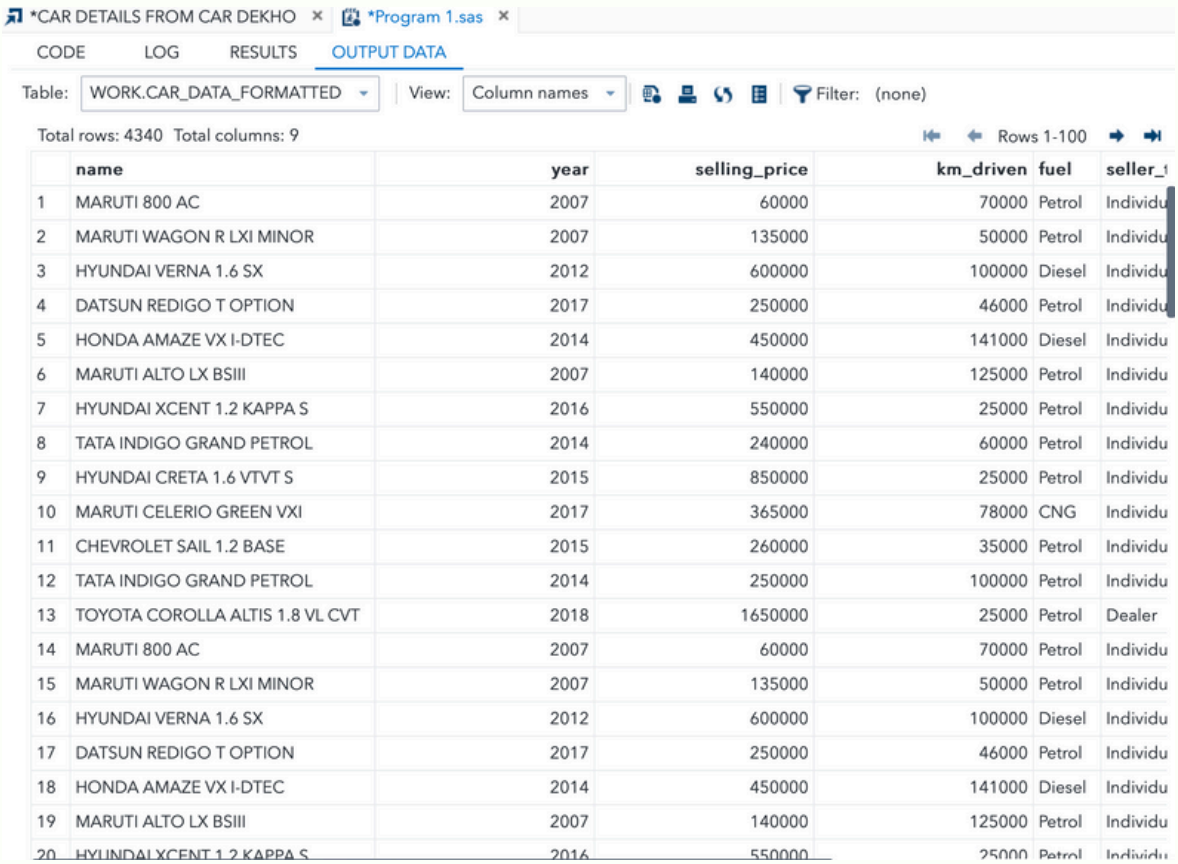
Car Dataset Using DOLLAR Format

Obs	name	year	selling_price	km_driven	fuel	seller_type	transmission	owner
1	Maruti 800 AC	2007	$60,000.00	70000	Petrol	Individual	Manual	First Owner
2	Maruti Wagon R LXI Minor	2007	$135,000.00	50000	Petrol	Individual	Manual	First Owner
3	Hyundai Verna 1.6 SX	2012	$600,000.00	100000	Diesel	Individual	Manual	First Owner
4	Datsun RediGO T Option	2017	$250,000.00	46000	Petrol	Individual	Manual	First Owner
5	Honda Amaze VX i-DTEC	2014	$450,000.00	141000	Diesel	Individual	Manual	Second Owner
6	Maruti Alto LX BSIII	2007	$140,000.00	125000	Petrol	Individual	Manual	First Owner
7	Hyundai Xcent 1.2 Kappa S	2016	$550,000.00	25000	Petrol	Individual	Manual	First Owner
8	Tata Indigo Grand Petrol	2014	$240,000.00	60000	Petrol	Individual	Manual	Second Owner
9	Hyundai Creta 1.6 VTVT S	2015	$850,000.00	25000	Petrol	Individual	Manual	First Owner
10	Maruti Celerio Green VXI	2017	$365,000.00	78000	CNG	Individual	Manual	First Owner
11	Chevrolet Sail 1.2 Base	2015	$260,000.00	35000	Petrol	Individual	Manual	First Owner
12	Tata Indigo Grand Petrol	2014	$250,000.00	100000	Petrol	Individual	Manual	First Owner
13	Toyota Corolla Altis 1.8 VL CVT	2018	$1650000.00	25000	Petrol	Dealer	Automatic	First Owner
14	Maruti 800 AC	2007	$60,000.00	70000	Petrol	Individual	Manual	First Owner
15	Maruti Wagon R LXI Minor	2007	$135,000.00	50000	Petrol	Individual	Manual	First Owner
16	Hyundai Verna 1.6 SX	2012	$600,000.00	100000	Diesel	Individual	Manual	First Owner
17	Datsun RediGO T Option	2017	$250,000.00	46000	Petrol	Individual	Manual	First Owner
18	Honda Amaze VX i-DTEC	2014	$450,000.00	141000	Diesel	Individual	Manual	Second Owner
19	Maruti Alto LX BSIII	2007	$140,000.00	125000	Petrol	Individual	Manual	First Owner
20	Hyundai Xcent 1.2 Kappa S	2016	$550,000.00	25000	Petrol	Individual	Manual	First Owner
21	Tata Indigo Grand Petrol	2014	$240,000.00	60000	Petrol	Individual	Manual	Second Owner
22	Hyundai Creta 1.6 VTVT S	2015	$850,000.00	25000	Petrol	Individual	Manual	First Owner
23	Maruti Celerio Green VXI	2017	$365,000.00	78000	CNG	Individual	Manual	First Owner
24	Chevrolet Sail 1.2 Base	2015	$260,000.00	35000	Petrol	Individual	Manual	First Owner

06

Code-Output

```
proc format;
 value fuel 1='Petrol' 2='Diesel' 3='CNG';
run;
data car_data_formatted;
 set car_data;
 formatted_fuel=put(fuel, fuel.);
run;
proc print data=car_data_formatted;
 title 'Formatted Car Dataset';
run;
```



*CAR DETAILS FROM CAR DEKHO × *Program 1.sas ×

CODE LOG RESULTS OUTPUT DATA

Table: WORK.CAR_DATA_FORMATTED View: Column names Filter: (none)

Total rows: 4340 Total columns: 9 Rows 1-100

	name	year	selling_price	km_driven	fuel	seller_t
1	MARUTI 800 AC	2007	60000	70000	Petrol	Individu
2	MARUTI WAGON R LXI MINOR	2007	135000	50000	Petrol	Individu
3	HYUNDAI VERNA 1.6 SX	2012	600000	100000	Diesel	Individu
4	DATSUN REDIGO T OPTION	2017	250000	46000	Petrol	Individu
5	HONDA AMAZE VX I-DTEC	2014	450000	141000	Diesel	Individu
6	MARUTI ALTO LX BSIII	2007	140000	125000	Petrol	Individu
7	HYUNDAI XCENT 1.2 KAPPA S	2016	550000	25000	Petrol	Individu
8	TATA INDIGO GRAND PETROL	2014	240000	60000	Petrol	Individu
9	HYUNDAI CRETA 1.6 VTVT S	2015	850000	25000	Petrol	Individu
10	MARUTI CELERIO GREEN VXI	2017	365000	78000	CNG	Individu
11	CHEVROLET SAIL 1.2 BASE	2015	260000	35000	Petrol	Individu
12	TATA INDIGO GRAND PETROL	2014	250000	100000	Petrol	Individu
13	TOYOTA COROLLA ALTIS 1.8 VL CVT	2018	1650000	25000	Petrol	Dealer
14	MARUTI 800 AC	2007	60000	70000	Petrol	Individu
15	MARUTI WAGON R LXI MINOR	2007	135000	50000	Petrol	Individu
16	HYUNDAI VERNA 1.6 SX	2012	600000	100000	Diesel	Individu
17	DATSUN REDIGO T OPTION	2017	250000	46000	Petrol	Individu
18	HONDA AMAZE VX I-DTEC	2014	450000	141000	Diesel	Individu
19	MARUTI ALTO LX BSIII	2007	140000	125000	Petrol	Individu
20	HYUNDAI XCENT 1.2 KAPPA S	2016	550000	25000	Petrol	Individu

07

Code-Output

```
DATA PAIRS;
 INPUT @1 ID 3.
 @6 (QN1-QN5) (1. +3) @7 (QC1-QC5) ($1. +3) @26 (HEIGHT AGE)
 (2. +1 2.);
 DATALINES;
187 1B 3B 4A 4A 6B 88 62
170 1B 3B 2B 2A 2B 98 31;

DATA random;
 INPUT @20 TYPE $1. @;
 IF TYPE='1' THEN
  INPUT ID 1-3 AGE 4-5 WEIGHT 6-8;
 ELSE IF TYPE='2' THEN
  INPUT ID 1-3 AGE 10-11 WEIGHT 15-17;
 DATALINES;
12345678          2
87654321          2
008       22  550  1;
proc univariate data=car_data;
 var selling_price;
 id name;
 title 'Summary of Car Data';
run;
```

Summary of Car Data

The UNIVARIATE Procedure
Variable: selling_price

Moments			
N	4340	Sum Weights	4340
Mean	504127.312	Sum Observations	2187912533
Std Deviation	578548.736	Variance	3.34719E11
Skewness	4.89202147	Kurtosis	37.0875426
Uncorrected SS	2.55533E15	Corrected SS	1.45234E15
Coeff Variation	114.762427	Std Error Mean	8782.0334

Basic Statistical Measures			
Location		Variability	
Mean	504127.3	Std Deviation	578549
Median	350000.0	Variance	3.34719E11
Mode	300000.0	Range	8880000
		Interquartile Range	392501

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	57.40439	Pr > |t|	<.0001
Sign	M	2170	Pr >= |M|	<.0001
Signed Rank	S	4709985	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	8900000
99%	3200000
95%	1300000
90%	900000
75% Q3	600000
50% Median	350000
25% Q1	207500
10%	110500
5%	80000
1%	55000
0% Min	20000

Extreme Observations					
Lowest			Highest		
Value	name	Obs	Value	name	Obs
20000	FORD IKON 1.6 ZXI NXT	2663	4950000	BMW X5 XDRIVE 30D XLINE	2259
22000	FORD IKON 1.4 ZXI	2496	4950000	BMW X5 XDRIVE 30D XLINE	2740
30000	MARUTI 800 EX	2445	5500000	MERCEDES-BENZ GLS 2016-2020 350	3970
35000	OPELCORSA 1.4 GL	3207	8150000	MERCEDES-BENZ S-CLASS S 350D CO	90
35000	TATA NANO STD BSII	2850	8900000	AUDI RS7 2015-2019 SPORTBACK PE	3873

Code-Output

```
proc univariate data=car_data nextrval=5 nextrobs=0;
 var selling_price;
 id name;
 title 'Summary of Car Data2';
run;
```

Summary of Car Data2

The UNIVARIATE Procedure
Variable: selling_price

Moments			
N	4340	Sum Weights	4340
Mean	504127.312	Sum Observations	2187912533
Std Deviation	578548.736	Variance	3.34719E11
Skewness	4.89202147	Kurtosis	37.0875426
Uncorrected SS	2.55533E15	Corrected SS	1.45234E15
Coeff Variation	114.762427	Std Error Mean	8782.0334

Basic Statistical Measures			
Location		Variability	
Mean	504127.3	Std Deviation	578549
Median	350000.0	Variance	3.34719E11
Mode	300000.0	Range	8880000
		Interquartile Range	392501

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	57.40439	Pr > |t|	<.0001
Sign	M	2170	Pr >= |M|	<.0001
Signed Rank	S	4709985	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	8900000
99%	3200000
95%	1300000
90%	900000
75% Q3	600000
50% Median	350000
25% Q1	207500
10%	110500
5%	80000
1%	55000
0% Min	20000

Extreme Values					
Lowest			Highest		
Order	Value	Freq	Order	Value	Freq
1	20000	1	441	4800000	2
2	22000	1	442	4950000	11
3	30000	1	443	5500000	1
4	35000	3	444	8150000	1
5	37500	1	445	8900000	1

09

Code-Output

```
proc means data=car_data;
 var selling_price;
 title 'Single-line Summary of Car Data';
run;
```

Single-line Summary of Car Data

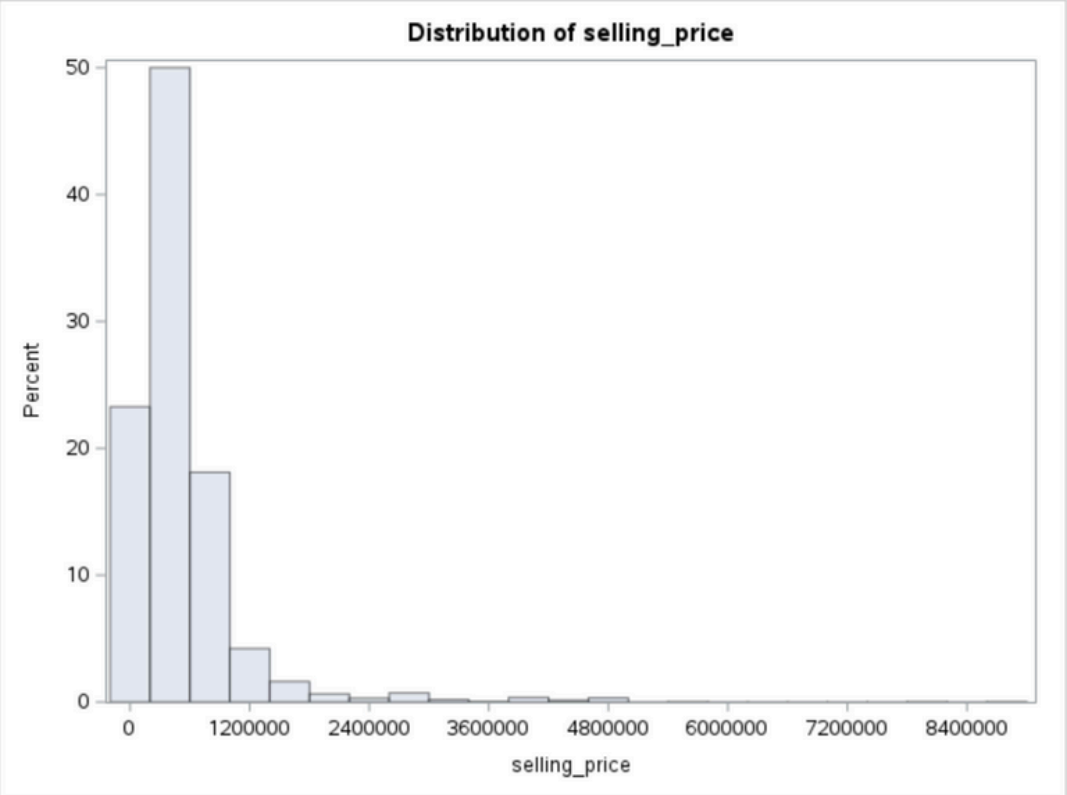
The MEANS Procedure

Analysis Variable : selling_price				
N	Mean	Std Dev	Minimum	Maximum
4340	504127.31	578548.74	20000.00	8900000.00

```
proc univariate data=car_data noprint;
 var selling_price;
 histogram selling_price;
 title 'Histogram for Car Data';
run;
```

Histogram for Car Data

The UNIVARIATE Procedure



Code-Output

```
proc univariate data=car_data freq;
 var year;
 title 'Summary of Cars Manufactures';
```

Summary of Cars Manufactures

The UNIVARIATE Procedure
Variable: year

Moments			
N	4340	Sum Weights	4340
Mean	2013.09078	Sum Observations	8736814
Std Deviation	4.21534394	Variance	17.7691245
Skewness	-0.8332399	Kurtosis	0.66826328
Uncorrected SS	1.75881E10	Corrected SS	77100.2313
Coeff Variation	0.20939661	Std Error Mean	0.06398647

Basic Statistical Measures			
Location		Variability	
Mean	2013.091	Std Deviation	4.21534
Median	2014.000	Variance	17.76912
Mode	2017.000	Range	28.00000
		Interquartile Range	5.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	31461.19	Pr > |t|	<.0001
Sign	M	2170	Pr >= |M|	<.0001
Signed Rank	S	4709985	Pr >= |S|	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	2020
99%	2020
95%	2019
90%	2018
75% Q3	2016
50% Median	2014
25% Q1	2011
10%	2007
5%	2005
1%	2001
0% Min	1992

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1992	3335	2020	3432
1995	632	2020	3487
1996	2973	2020	3934
1996	62	2020	4106
1997	3662	2020	4279

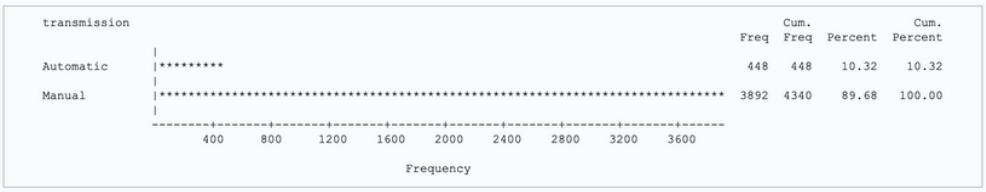
Frequency Counts			
		Percents	
Value	Count	Cell	Cum
1992	1	0.0	0.0
1995	1	0.0	0.0
1996	2	0.0	0.1
1997	3	0.1	0.2
1998	12	0.3	0.4
1999	10	0.2	0.7
2000	12	0.3	0.9
2001	20	0.5	1.4
2002	21	0.5	1.9
2003	23	0.5	2.4
2004	42	1.0	3.4
2005	85	2.0	5.3
2006	110	2.5	7.9
2007	134	3.1	11.0
2008	145	3.3	14.3
2009	193	4.4	18.8
2010	234	5.4	24.1
2011	271	6.2	30.4
2012	415	9.6	40.0
2013	386	8.9	48.8
2014	367	8.5	57.3
2015	421	9.7	67.0
2016	357	8.2	75.2
2017	466	10.7	86.0
2018	366	8.4	94.4
2019	195	4.5	98.9
2020	48	1.1	100.0

Code-Output

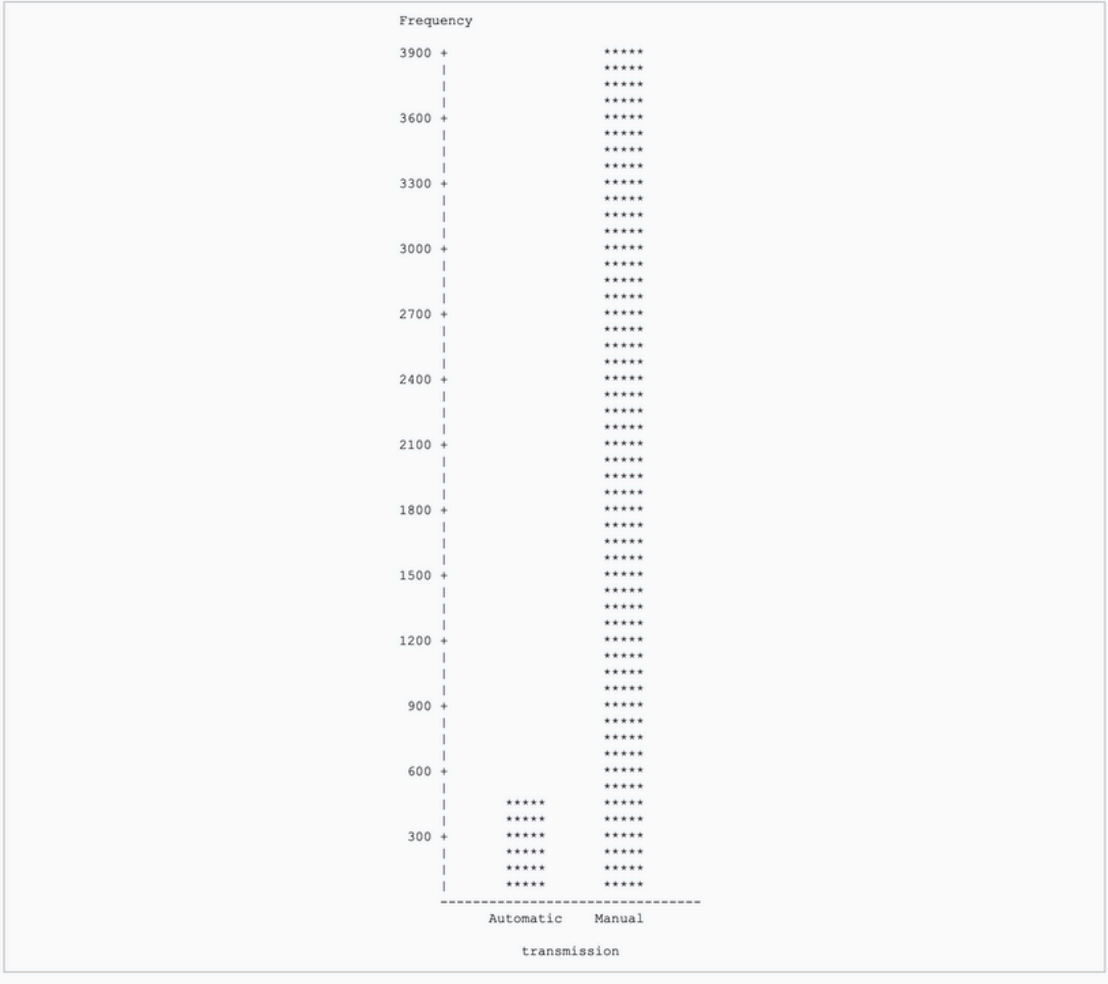
```
proc freq data=car_data;
 title 'Frequance Table';
run;
goptions device=win;
pattern v=solid color=gray;
proc chart data=car_data;
 vbar transmission;
 title 'Bar Chart for Cars Transmission Type';
run;
proc chart data=car_data;
 hbar transmission;
 title 'Horizontal Bar Chart for Cars Transmission Type';
run;
```

Horizontal Bar Chart for Cars Transmission Type

transmission	Freq	Cum. Freq	Percent	Cum. Percent
Automatic	448	448	10.32	10.32
Manual	3892	4340	89.68	100.00



Bar Chart for Cars Transmission Type



Code-Output

```
proc means data=car_data n mean stddev clm alpha=0.10;
 var km_driven;
 title 'Summary of Km Driven Data with 90% CI';
run;
data car_data;
 car_age=2024-year;
 label car_age='Vehicle Age';
run;
```

Summary of Km Driven Data with 90% CI

The MEANS Procedure

Analysis Variable : km_driven Kilometers Driven				
N	Mean	Std Dev	Lower 90% CL for Mean	Upper 90% CL for Mean
4340	66215.78	46644.10	65050.92	67380.63

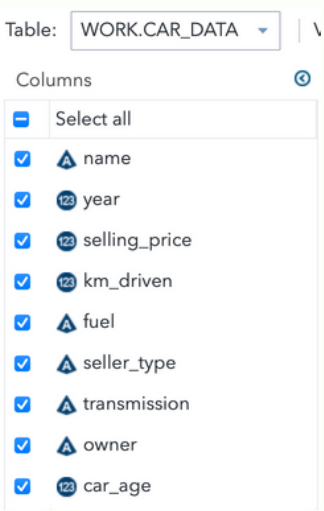
Single-line Summary of Car Data

The MEANS Procedure

Analysis Variable : selling_price Selling Price								
N	N Miss	Mean	Median	Std Dev	Range	Quartile Range	Minimum	Maximum
4340	0	504127.31	350000.00	578548.74	8880000.00	392500.50	20000.00	8900000.00

```
data car_data;
 car_age=2024-year;
 label car_age='Vehicle Age';
run;
ods select TestsForLocation;
```

With this code we create a new varible named 'car_age'. it contains vehicle age



Code-Output

```
proc ttest data=car_data;
 paired selling_price*km_driven;
 title 'Paired Differences with PROC TTEST';
run;
proc chart data=car_data;
 vbar year / group=fuel;
 title 'Charts for Fuel Usage';
run;
```

Paired Differences with PROC TTEST

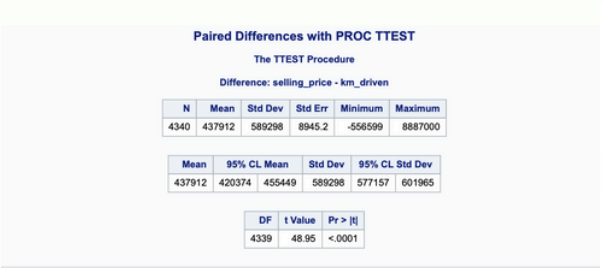
The TTEST Procedure

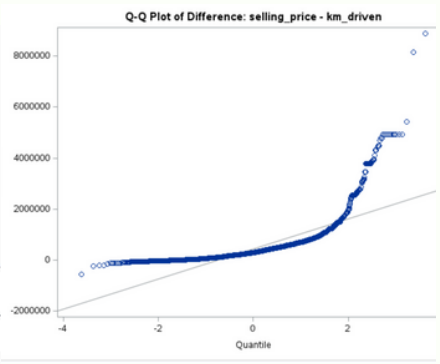
Difference: selling_price - km_driven

N	Mean	Std Dev	Std Err	Minimum	Maximum
4340	437912	589298	8945.2	-556599	8887000

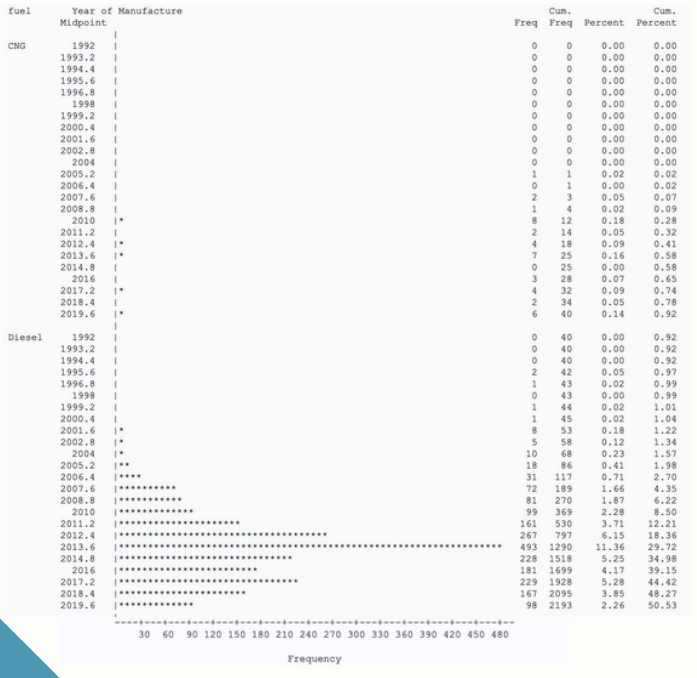
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
437912	420374	455449	589298	577157	601965

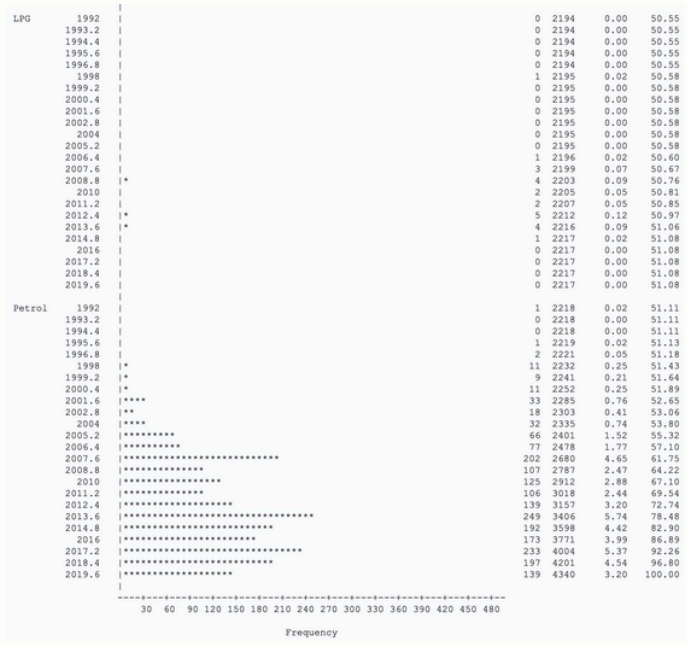
DF	t Value	Pr > |t|
4339	48.95	<.0001





Charts for Fuel Usage





Code-Output

```
proc npar1way data=car_data wilcoxon;
 class fuel;
 var year;
 title 'Comparison of Ulcer and Control Patients';
run;
proc gchart data=car_data;
 hbar transmission / nostat descending sumvar=km_driven;
 run;
```

Comparison of Ulcer and Control Patients

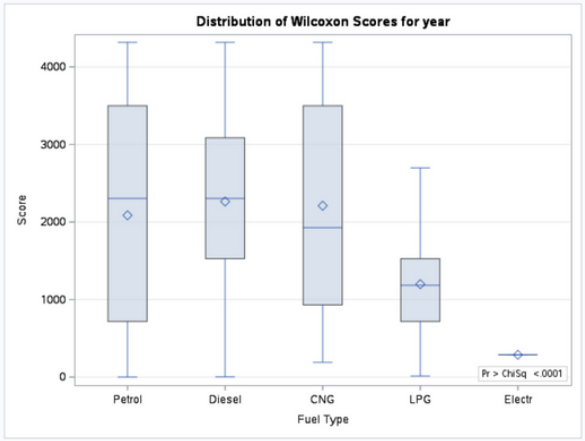
The NPAR1WAY Procedure

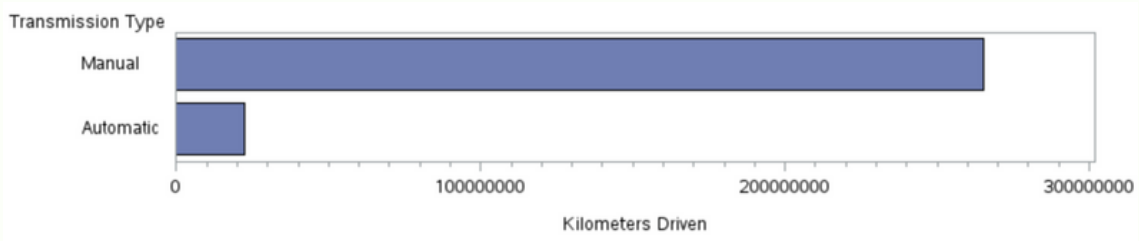
Wilcoxon Scores (Rank Sums) for Variable year Classified by Variable fuel

fuel	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Petrol	2123	4428993.50	4607971.50	41135.8046	2086.19571
Diesel	2153	4874723.50	4673086.50	41144.1940	2264.15397
CNG	40	88397.50	86820.00	7863.6899	2209.93750
LPG	23	27568.00	49921.50	5974.7136	1198.60870
Electr	1	287.50	2170.50	1248.9843	287.50000

Average scores were used for ties.

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
38.0087	4	<.0001

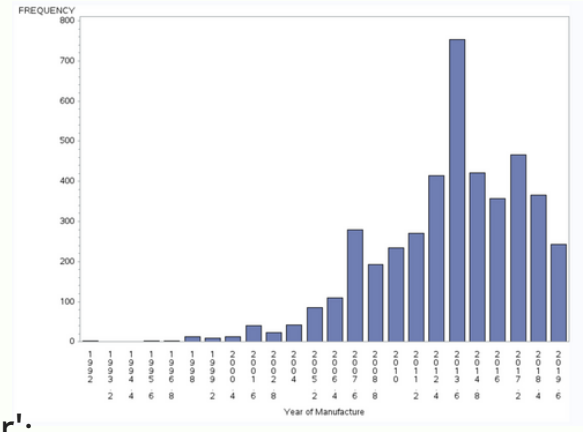




Code-Output

```
proc gchart data=car_data normal;
 var year;
 run;

proc univariate data=car_data;
 var selling_price year;
 histogram selling_price / normal(color=red);
 histogram year / normal(color=blue);
 title 'Normality Test for Selling Price and Year';
run;
```



Normality Test for Selling Price and Year

The UNIVARIATE Procedure
Variable: selling_price (Selling Price)

Moments			
N	4340	Sum Weights	4340
Mean	504127.312	Sum Observations	2187912533
Std Deviation	578548.736	Variance	3.34719E11
Skewness	4.89202147	Kurtosis	37.0875426
Uncorrected SS	2.55533E15	Corrected SS	1.45234E15
Coeff Variation	114.762427	Std Error Mean	8782.0334

Basic Statistical Measures			
Location		Variability	
Mean	504127.3	Std Deviation	578549
Median	350000.0	Variance	3.34719E11
Mode	300000.0	Range	8880000
		Interquartile Range	392501

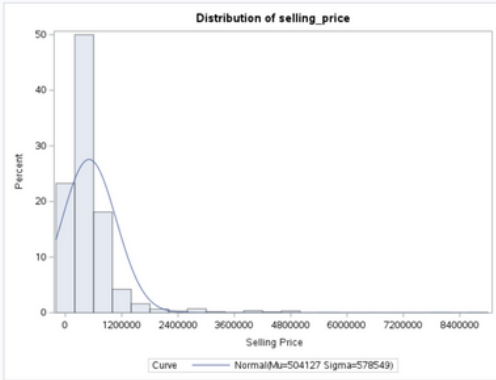
Tests for Location: Mu0=0			
Test	Statistic		p Value
Student's t	t	57.40439	Pr > |t| <.0001
Sign	M	2170	Pr >= |M| <.0001
Signed Rank	S	4709985	Pr >= |S| <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	8900000
99%	3200000
95%	1300000
90%	900000
75% Q3	600000
50% Median	350000
25% Q1	207500
10%	110500
5%	80000
1%	55000
0% Min	20000

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
20000	2663	4950000	2259
22000	2496	4950000	2740
30000	2445	5500000	3970
35000	3207	8150000	90
35000	2850	8900000	3873

Normality Test for Selling Price and Year

The UNIVARIATE Procedure



Normality Test for Selling Price and Year

The UNIVARIATE Procedure
Fitted Normal Distribution for selling_price (Selling Price)

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	504127.3
Std Dev	Sigma	578548.7

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.210483	Pr > D	<0.010
Cramer-von Mises	W-Sq	70.095001	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	396.418783	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
	Quantile	
Percent	Observed	Estimated
1.0	55000.0	-841778
5.0	80000.0	-447501
10.0	110500.0	-237313
25.0	207499.5	113902
50.0	350000.0	504127
75.0	600000.0	894353
90.0	900000.0	1245567
95.0	1300000.0	1455755
99.0	3200000.0	1850033

Normality Test for Selling Price and Year

The UNIVARIATE Procedure
Variable: year (Year of Manufacture)

Moments			
N	4340	Sum Weights	4340
Mean	2013.09078	Sum Observations	8736814
Std Deviation	4.21534394	Variance	17.7691245
Skewness	-0.8332399	Kurtosis	0.66826328
Uncorrected SS	1.75881E10	Corrected SS	77100.2313
Coeff Variation	0.20939661	Std Error Mean	0.06398647

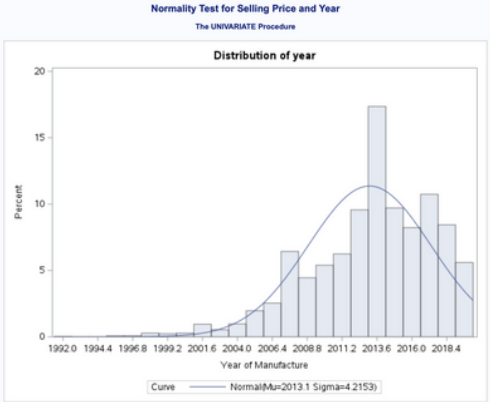
Basic Statistical Measures			
Location		Variability	
Mean	2013.091	Std Deviation	4.21534
Median	2014.000	Variance	17.76912
Mode	2017.000	Range	28.00000
		Interquartile Range	5.00000

Tests for Location: Mu0=0			
Test	Statistic		p Value
Student's t	t	31461.19	Pr > |t| <.0001
Sign	M	2170	Pr >= |M| <.0001
Signed Rank	S	4709985	Pr >= |S| <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	2020
99%	2020
95%	2019
90%	2018
75% Q3	2016
50% Median	2014
25% Q1	2011
10%	2007
5%	2005
1%	2001
0% Min	1992

Normality Test for Selling Price and Year

The UNIVARIATE Procedure



Normality Test for Selling Price and Year

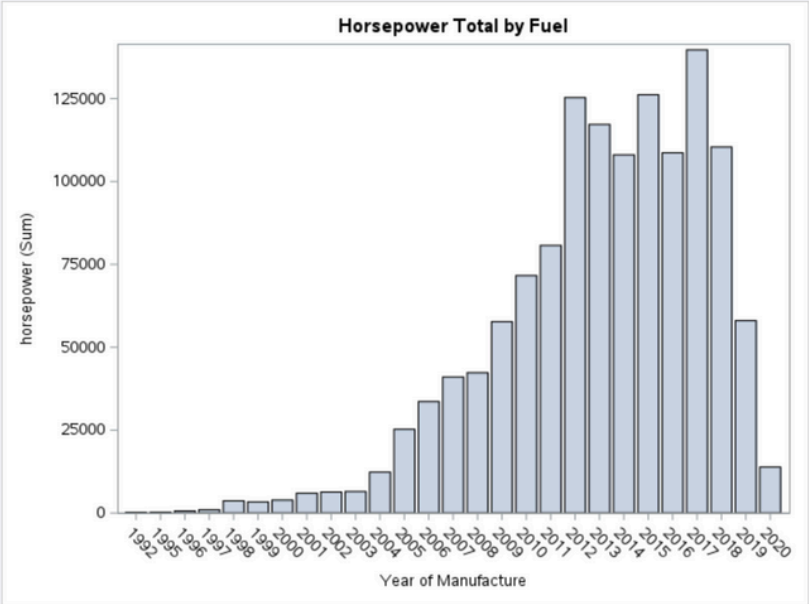
The UNIVARIATE Procedure
Fitted Normal Distribution for year (Year of Manufacture)

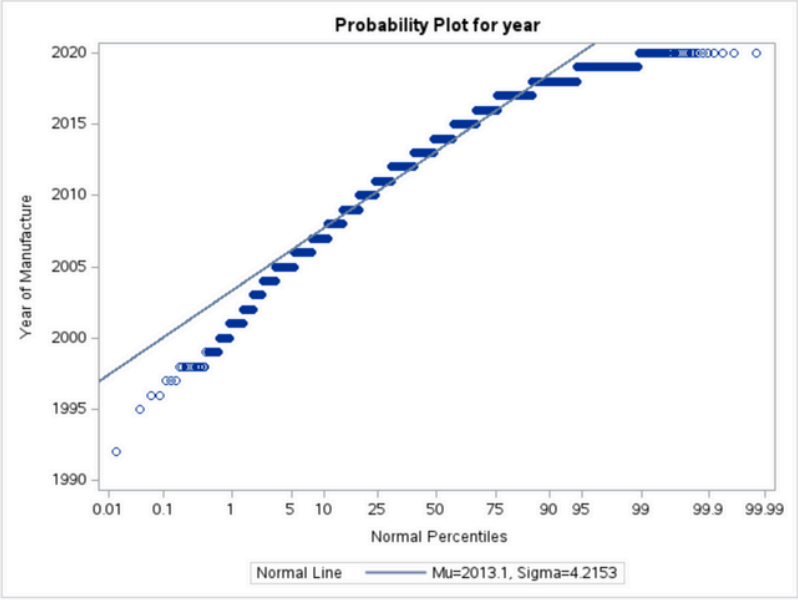
Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	2013.091
Std Dev	Sigma	4.215344

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.1016556	Pr > D	<0.010
Cramer-von Mises	W-Sq	8.4294830	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	54.7003393	Pr > A-Sq	<0.005

Code-Output

```
proc sgplot data=car_data;
 vbar year / response=horsepower stat=sum;
 title 'Horsepower Total Over Years';
run;
proc univariate data=car_data;
 var year;
 probplot year / normal(mu=est sigma=est color=red);
run;
```





The UNIVARIATE Procedure
Variable: year (Year of Manufacture)

Moments			
N	4340	Sum Weights	4340
Mean	2013.09078	Sum Observations	8736814
Std Deviation	4.21534394	Variance	17.7691245
Skewness	-0.8332399	Kurtosis	0.66826328
Uncorrected SS	1.75881E10	Corrected SS	77100.2313
Coeff Variation	0.20939661	Std Error Mean	0.06398647

Basic Statistical Measures			
Location		Variability	
Mean	2013.091	Std Deviation	4.21534
Median	2014.000	Variance	17.76912
Mode	2017.000	Range	28.00000
		Interquartile Range	5.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	31461.19	Pr > |t|	<.0001
Sign	M	2170	Pr >= |M|	<.0001
Signed Rank	S	4709985	Pr >= |S|	<.0001

Code-Output

```
proc means data=car_data n mean stddev clm maxdec=2;
 var horsepower;
run;
```

The MEANS Procedure

Analysis Variable : horsepower				
N	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
4340	300.43	57.14	298.73	302.13

```
data bodyfat;
 input gender $ fatpct @@;
 format gender $gentext.;
 label fatpct='Body Fat Percentage';
 datalines;
f 18.7 m 26 f 15 m 21 f 16 m 21 f 6 m 9 f 13 m 25.2
f 20 m 23.2 f 17 m 21 f 29 m 32 f 11 m 20 f 9 m 27
f 15 f 11 f 20
;
run;
proc means data=bodyfat;
 class gender;
 var fatpct;
 title 'Brief Summary of Groups';
run;
ods select moments basicmeasures extremeobs plots;
```

Brief Summary of Groups

The MEANS Procedure

Analysis Variable : fatpct Body Fat Percentage						
gender	N Obs	N	Mean	Std Dev	Minimum	Maximum
f	13	13	15.4384615	5.8821111	6.0000000	29.0000000
m	10	10	22.5400000	6.0163112	9.0000000	32.0000000

Code-Output

```
proc univariate data=bodyfat plot;
 class gender;
 var fatpct;
 title 'Detailed Summary of Groups';
run;
```

Detailed Summary of Groups

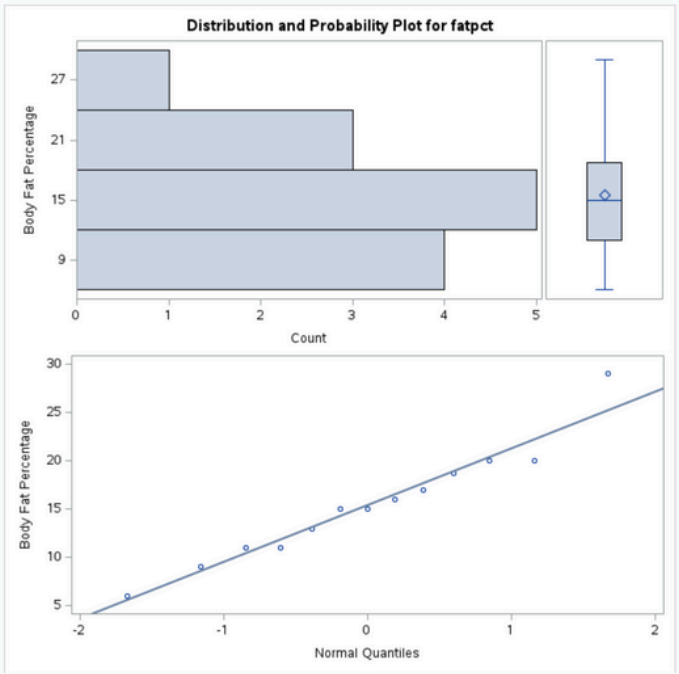
The UNIVARIATE Procedure
Variable: fatpct (Body Fat Percentage)
gender = f

Moments			
N	13	Sum Weights	13
Mean	15.4384615	Sum Observations	200.7
Std Deviation	5.88211108	Variance	34.5992308
Skewness	0.68978398	Kurtosis	1.28747216
Uncorrected SS	3513.69	Corrected SS	415.190769
Coeff Variation	38.1003707	Std Error Mean	1.63140409

Basic Statistical Measures			
Location		Variability	
Mean	15.43846	Std Deviation	5.88211
Median	15.00000	Variance	34.59923
Mode	11.00000	Range	23.00000
		Interquartile Range	7.70000

Note: The mode displayed is the smallest of 3 modes with a count of 2.

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
6	7	17.0	13
9	19	18.7	1
11	22	20.0	11
11	17	20.0	23
13	9	29.0	15



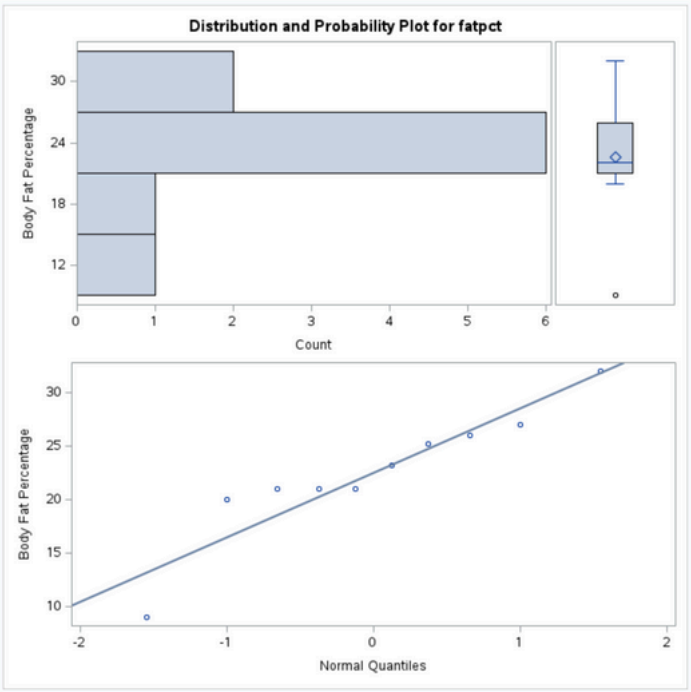
Detailed Summary of Groups

The UNIVARIATE Procedure
Variable: fatpct (Body Fat Percentage)
gender = m

Moments			
N	10	Sum Weights	10
Mean	22.54	Sum Observations	225.4
Std Deviation	6.01631116	Variance	36.196
Skewness	-0.9654945	Kurtosis	2.70183969
Uncorrected SS	5406.28	Corrected SS	325.764
Coeff Variation	26.6917088	Std Error Mean	1.90252464

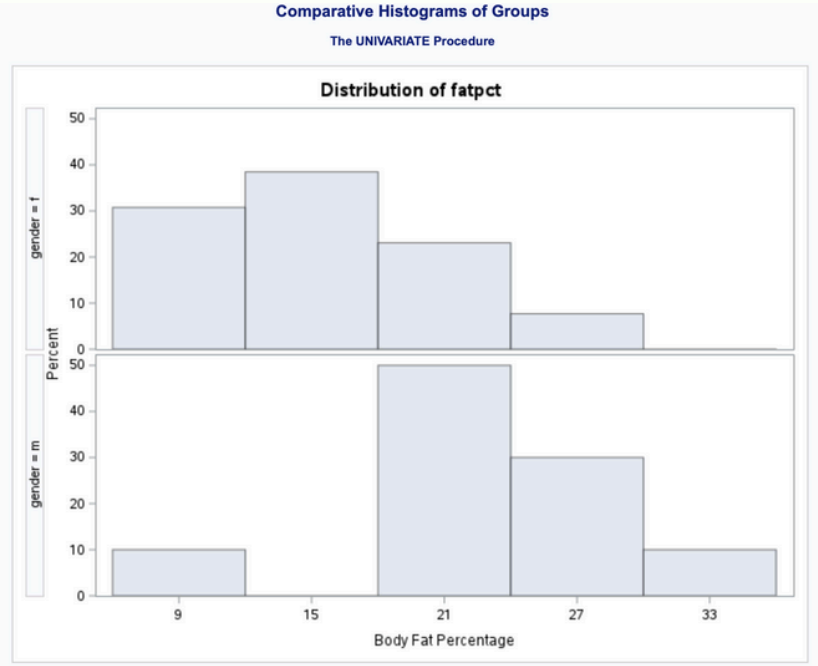
Basic Statistical Measures			
Location		Variability	
Mean	22.54000	Std Deviation	6.01631
Median	22.10000	Variance	36.19600
Mode	21.00000	Range	23.00000
		Interquartile Range	5.00000

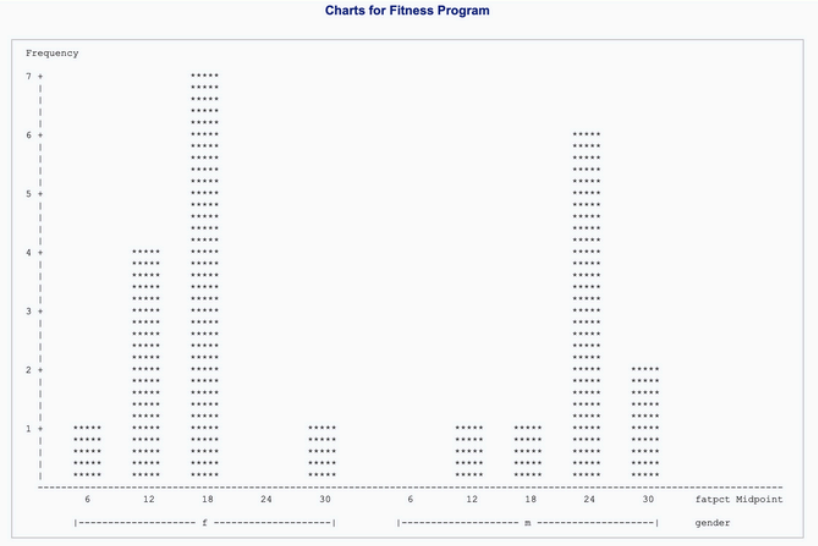
Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
9	8	23.2	12
20	18	25.2	10
21	14	26.0	2
21	6	27.0	20
21	4	32.0	16



Code-Output

```
proc univariate data=bodyfat noprint;
 class gender;
 var fatpct;
 histogram fatpct;
 title 'Comparative Histograms of Groups';
run;
proc chart data=bodyfat;
 vbar fatpct / group=gender;
 title 'Charts for Fitness Program';
run;
```





Code-Output

```
proc ttest data=bodyfat;
 class gender;
 var fatpct;
 title 'Comparing Groups in Fitness Program';
run;
```

Comparing Groups in Fitness Program

The TTEST Procedure

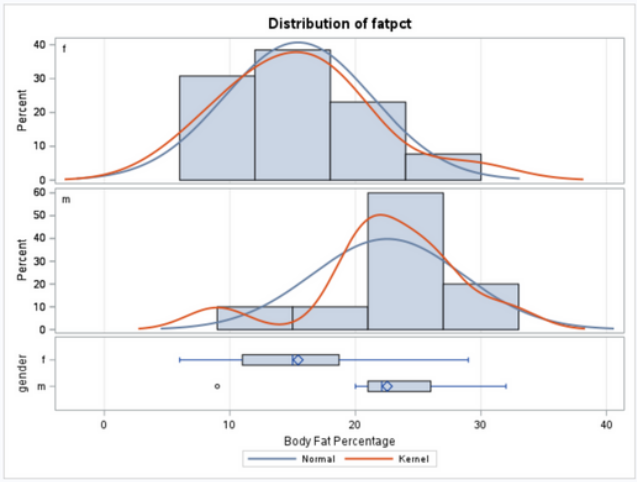
Variable: fatpct (Body Fat Percentage)

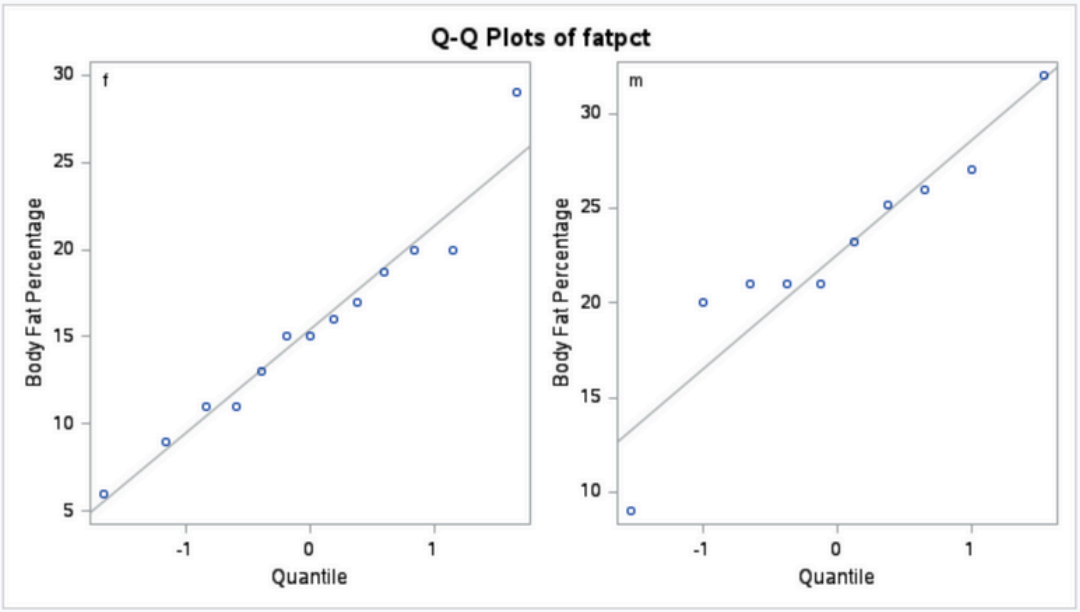
gender	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
f		13	15.4385	5.8821	1.6314	6.0000	29.0000
m		10	22.5400	6.0163	1.9025	9.0000	32.0000
Diff (1-2)	Pooled		-7.1015	5.9400	2.4985		
Diff (1-2)	Satterthwaite		-7.1015		2.5062		

gender	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
f		15.4385	11.8839	18.9930	5.8821	4.2180	9.7098
m		22.5400	18.2362	26.8438	6.0163	4.1382	10.9834
Diff (1-2)	Pooled	-7.1015	-12.2974	-1.9056	5.9400	4.5699	8.4886
Diff (1-2)	Satterthwaite	-7.1015	-12.3419	-1.8612			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	21	-2.84	0.0098
Satterthwaite	Unequal	19.282	-2.83	0.0105

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	9	12	1.05	0.9197





Code-Output

```
proc npar1way data=bodyfat wilcoxon;
 class gender;
 var fatpct;
 title 'Comparison of Gender';
run;
ods select wilcoxontest;
```

Comparison of Gender

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable fatpct Classified by Variable gender

gender	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
f	13	111.50	156.0	16.080639	8.576923
m	10	164.50	120.0	16.080639	16.450000

Average scores were used for ties.

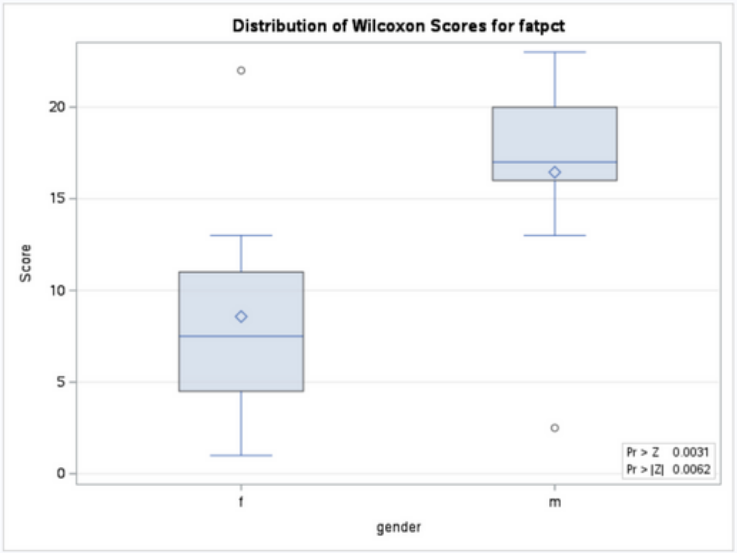
Wilcoxon Two-Sample Test

Statistic	Z	Pr > Z	Pr > |Z|	t Approximation Pr > Z	t Approximation Pr > |Z|
164.5000	2.7362	0.0031	0.0062	0.0060	0.0121

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	DF	Pr > ChiSq
7.6580	1	0.0057



Code-Output

```
proc npar1way data=bodyfat wilcoxon;
 class gender;
 var fatpct;
run;
```

The NPAR1WAY Procedure

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr > Z	Pr > |Z|	Pr > Z	Pr > |Z|
164.5000	2.7362	0.0031	0.0062	0.0060	0.0121
Z includes a continuity correction of 0.5.					

```
data data2;
 input gender $ age weight height;
 datalines;
M 25 70 185
F 30 65 150
F 35 80 152
F 28 55 172
M 40 85 166
;
run;
data data2;
 set data2;
 bmi=weight / (height * height) * 10000;
run;
proc means data=data2 n mean stddev clm maxdec=3;
 var bmi;
 title 'sumarry of body mass index';
run;
proc means data=data2 n mean stddev clm alpha=0.01;
 var bmi;
 title 'Summary of Body Mass index Data with 99% CI';
run;
proc means data=data2;
 class gender;
 var bmi;
 title 'Brief Summary of bmi';
run;
```

sumarry of body mass index

The MEANS Procedure

Analysis Variable : bmi				
N	Mean	Std Dev	Lower 95% CL for Mean	Upper 95% CL for Mean
5	26.681	6.884	18.133	35.229

Summary of Body Mass index Data with 99% CI

The MEANS Procedure

Analysis Variable : bmi				
N	Mean	Std Dev	Lower 99% CL for Mean	Upper 99% CL for Mean
5	26.6810440	6.8844634	12.5058386	40.8562494

Brief Summary of bmi

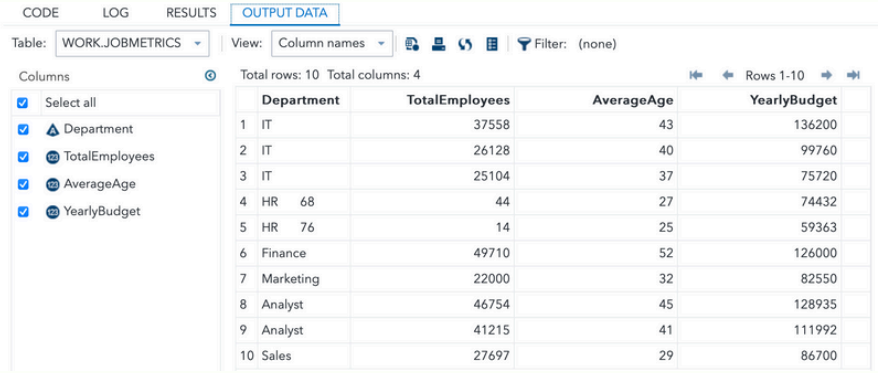
The MEANS Procedure

Analysis Variable : bmi						
gender	N Obs	N	Mean	Std Dev	Minimum	Maximum
F	3	3	27.3686860	8.1248283	18.5911303	34.6260388
M	2	2	25.6495810	7.3492375	20.4528853	30.8462767

	gender	age	weight	height	bmi
1	M	25	70	185	20.452885318
2	F	30	65	150	28.888888889
3	F	35	80	152	34.626038781
4	F	28	55	172	18.591130341
5	M	40	85	166	30.846276673

Code-Output

```
data jobmetrics;
 input Department $ 1-12 TotalEmployees AverageAge YearlyBudget;
 datalines;
IT          37558 43 136200
IT          26128 40 99760
IT          25104 37 75720
HR          6844 27 74432
HR          7614 25 59363
Finance       49710 52 126000
Marketing     22000 32 82550
Analyst       46754 45 128935
Analyst       41215 41 111992
Sales         27697 29 86700
;
run;
```



	Department	TotalEmployees	AverageAge	YearlyBudget
1	IT	37558	43	136200
2	IT	26128	40	99760
3	IT	25104	37	75720
4	HR 68	44	27	74432
5	HR 76	14	25	59363
6	Finance	49710	52	126000
7	Marketing	22000	32	82550
8	Analyst	46754	45	128935
9	Analyst	41215	41	111992
10	Sales	27697	29	86700

```
proc print data=jobmetrics;
run;
```

Obs	Department	TotalEmployees	AverageAge	YearlyBudget
1	IT	37558	43	136200
2	IT	26128	40	99760
3	IT	25104	37	75720
4	HR	6844	27	74432
5	HR	7614	25	59363
6	Finance	49710	52	126000
7	Marketing	22000	32	82550
8	Analyst	46754	45	128935
9	Analyst	41215	41	111992
10	Sales	27697	29	86700

Code-Output

```
proc anova data=jobmetrics;
 class Department;
 model TotalEmployees=Department;
 ;
 title 'ANOVA for Department Salaries';
 run;
```

ANOVA for Department Salaries

The ANOVA Procedure

Class Level Information		
Class	**Levels**	**Values**
Department	7	Analyst Analyst Finance HR IT Marketing Sales

Number of Observations Read	10
Number of Observations Used	10

ANOVA for Department Salaries

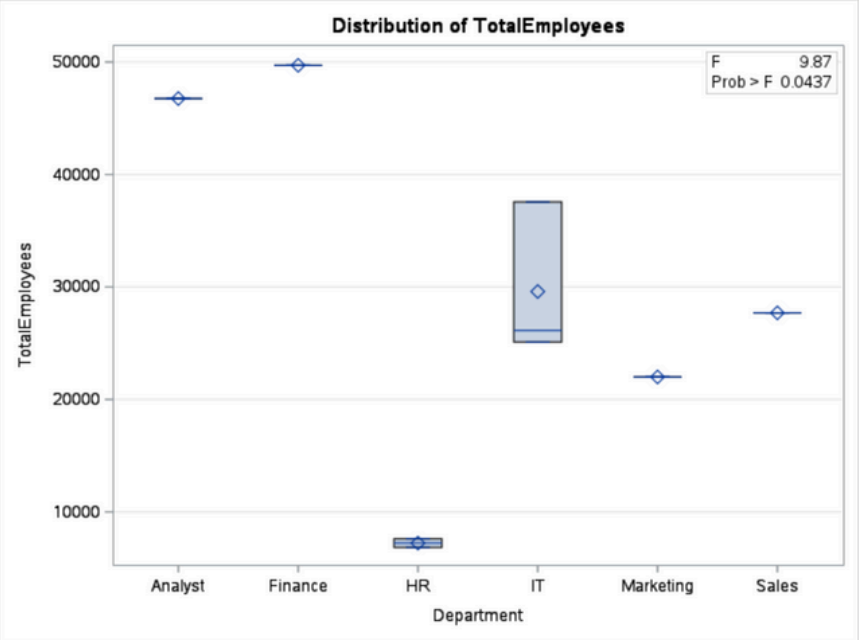
The ANOVA Procedure

Dependent Variable: TotalEmployees

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	1892994628	315499105	9.87	0.0437
Error	3	95894981	31964994		
Corrected Total	9	1988889608			

R-Square	Coeff Var	Root MSE	TotalEmployees Mean
0.951785	19.45386	5653.759	29062.40

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Department	6	1892994628	315499105	9.87	0.0437



Code-Output

```
proc npar1way data=jobmetrics wilcoxon;
 class Department;
 var YearlyBudget;
 title 'Nonparametric Tests for Job Metrics Data';
run;
```

Nonparametric Tests for Job Metrics Data

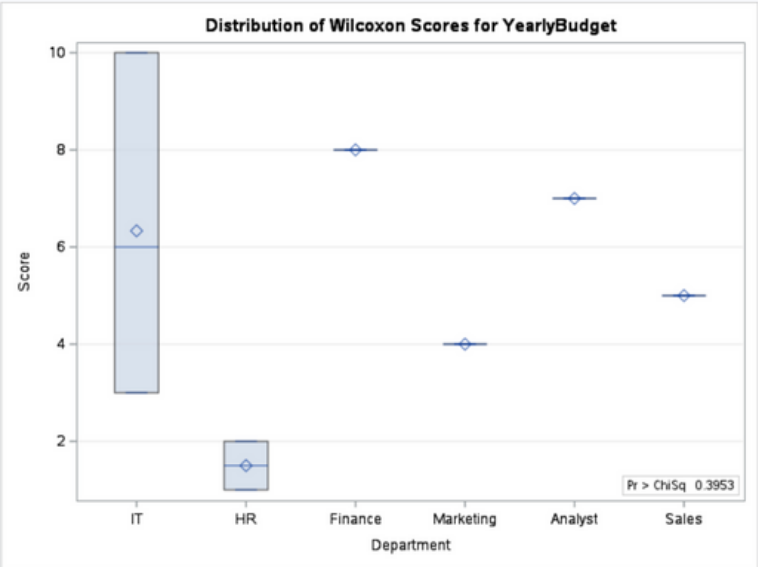
The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable YearlyBudget Classified by Variable Department

Department	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
IT	3	19.0	16.50	4.387482	6.333333
HR	2	3.0	11.00	3.829708	1.500000
Finance	1	8.0	5.50	2.872281	8.000000
Marketing	1	4.0	5.50	2.872281	4.000000
Analyst	1	9.0	5.50	2.872281	9.000000
Analyst	1	7.0	5.50	2.872281	7.000000
Sales	1	5.0	5.50	2.872281	5.000000

Kruskal-Wallis Test

Chi-Square	DF	Pr > ChiSq
6.2545	6	0.3953



Code-Output

```
proc anova data=jobmetrics;
 class Department;
 model YearlyBudget=Department;
 means Department / t;
 title 'Multiple Comparisons with t Tests';
 run;
```

Multiple Comparisons with t Tests

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Department	7	Analyst Analyst Finance HR IT Marketing Sales

Number of Observations Read	10
Number of Observations Used	10

Multiple Comparisons with t Tests

The ANOVA Procedure

Dependent Variable: YearlyBudget

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	4341795024	723632504	1.10	0.5086
Error	3	1968079247	656026416		
Corrected Total	9	6309874272			

R-Square	Coeff Var	Root MSE	YearlyBudget Mean
0.688095	26.09174	25613.01	98165.20

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Department	6	4341795024	723632504	1.10	0.5086

Multiple Comparisons with t Tests

The ANOVA Procedure

t Tests (LSD) for YearlyBudget

Note: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	3
Error Mean Square	6.5603E8
Critical Value of t	3.18245

Comparisons significant at the 0.05 level are indicated by ***.

Department Comparison	Difference Between Means	95% Confidence Limits	
Analyst - Finance	2935	-112340	118210
Analyst - Analyst	16943	-98332	132218
Analyst - IT	25042	-69080	119164
Analyst - Sales	42235	-73040	157510
Analyst - Marketing	46385	-68890	161660
Analyst - HR	62038	-37794	161869
Finance - Analyst	-2935	-118210	112340
Finance - Analyst	14008	-101267	129283
Finance - IT	22107	-72015	116229
Finance - Sales	39300	-75975	154575
Finance - Marketing	43450	-71825	158725
Finance - HR	59103	-40729	158934
Analyst - Analyst	-16943	-132218	98332
Analyst - Finance	-14008	-129283	101267
Analyst - IT	8099	-86023	102221
Analyst - Sales	25292	-89983	140567
Analyst - Marketing	29442	-85833	144717
Analyst - HR	45095	-54737	144926
IT - Analyst	-25042	-119164	69080
IT - Finance	-22107	-116229	72015
IT - Analyst	-8099	-102221	86023
IT - Sales	17193	-76929	111315
IT - Marketing	21343	-72779	115465
IT - HR	36996	-37414	111406
Sales - Analyst	-42235	-157510	73040
Sales - Finance	-39300	-154575	75975
Sales - Analyst	-25292	-140567	89983
Sales - IT	-17193	-111315	76929
Sales - Marketing	4150	-111125	119425
Sales - HR	19803	-80029	119634
Marketing - Analyst	-46385	-161660	68890
Marketing - Finance	-43450	-158725	71825
Marketing - Analyst	-29442	-144717	85833
Marketing - IT	-21343	-115465	72779
Marketing - Sales	-4150	-119425	111125
Marketing - HR	15653	-84179	115484
HR - Analyst	-62038	-161869	37794
HR - Finance	-59103	-158934	40729
HR - Analyst	-45095	-144926	54737
HR - IT	-36996	-111406	37414
HR - Sales	-19803	-119634	80029
HR - Marketing	-15653	-115484	84179

Code-Output

```
proc corr data=jobmetrics plots=scatter(noinset ellipse=none);
 var YearlyBudget AverageAge;
run;
ods graphics off;
ods graphics on;
ods select MatrixPlot;
proc corr data=jobmetrics;
 var YearlyBudget AverageAge TotalEmployees;
 title 'Summary Statistics for Job Parametrics Data Set';
run;
```

The CORR Procedure

2 Variables:	YearlyBudget AverageAge

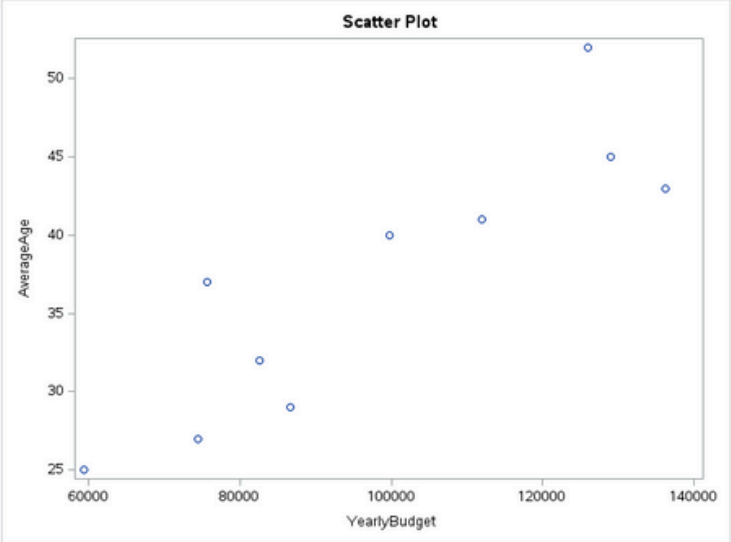
Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
YearlyBudget	10	98165	26478	981652	59363	136200
AverageAge	10	37.10000	8.71079	371.00000	25.00000	52.00000

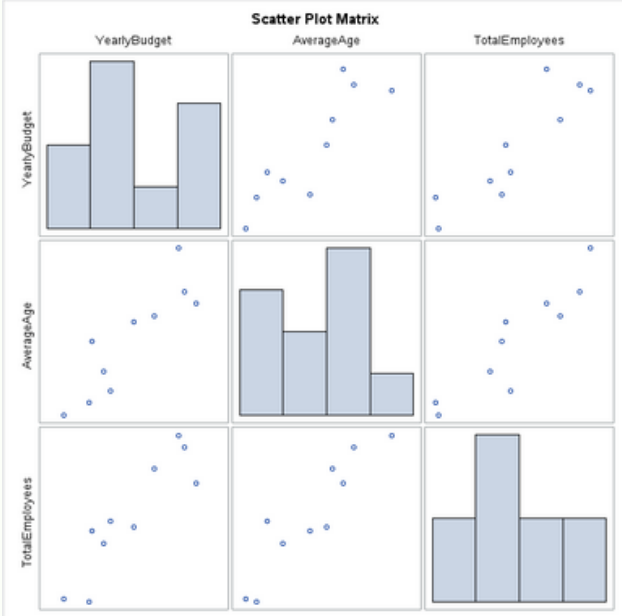
Pearson Correlation Coefficients, N = 10
Prob > |r| under H0: Rho=0

	YearlyBudget	AverageAge
YearlyBudget	1.00000	0.87905 0.0008
AverageAge	0.87905 0.0008	1.00000

The CORR Procedure



The CORR Procedure



Summary Statistics for Job Parametrics Data Set

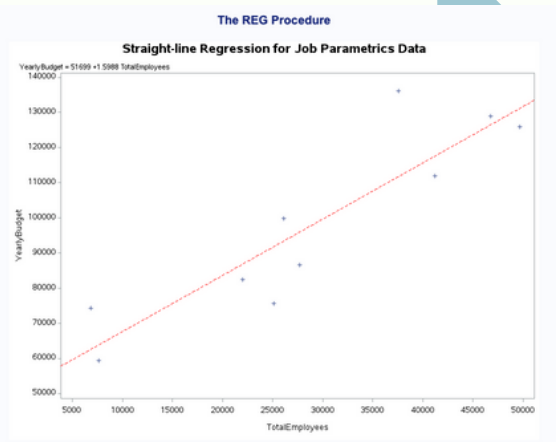
The CORR Procedure

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
YearlyBudget	10	98165	26478	981652	59363	136200
AverageAge	10	37.10000	8.71079	371.00000	25.00000	52.00000
TotalEmployees	10	29062	14866	290624	6844	49710

Code-Output

```
proc reg data=jobmetrics;
 model YearlyBudget=TotalEmployees;
 plot YearlyBudget*TotalEmployees / nostat cline=red;
 title 'Straight-line Regression for Job Parametrics Data';
 run;
 ods select SimpleStats;
 ods graphics on
proc reg data=car_data plots(only)=fit(stats=none);
 model selling_price=km_driven;
 title 'Regression for Car Data';
 run;
quit;
```



Straight-line Regression for Job Parametrics Data

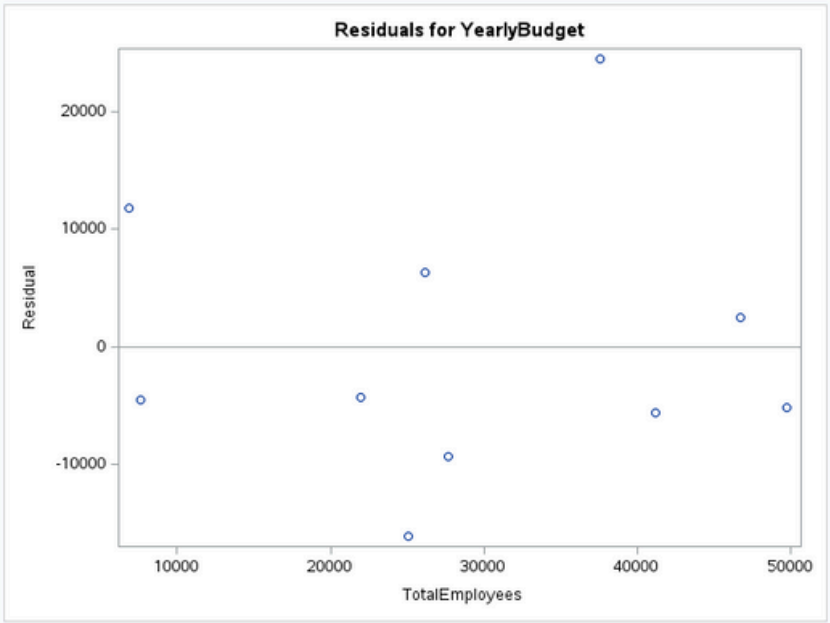
The REG Procedure
Model: MODEL1
Dependent Variable: YearlyBudget

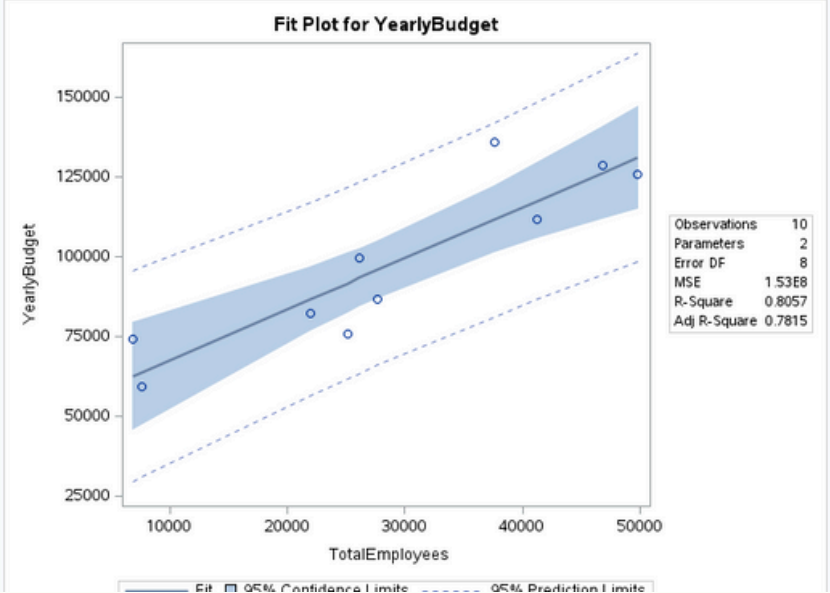
Number of Observations Read	10
Number of Observations Used	10

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	5084158365	5084158365	33.18	0.0004
Error	8	1225715906	153214488		
Corrected Total	9	6309874272			

Root MSE	12378	R-Square	0.8057
Dependent Mean	98165	Adj R-Sq	0.7815
Coeff Var	12.60934		

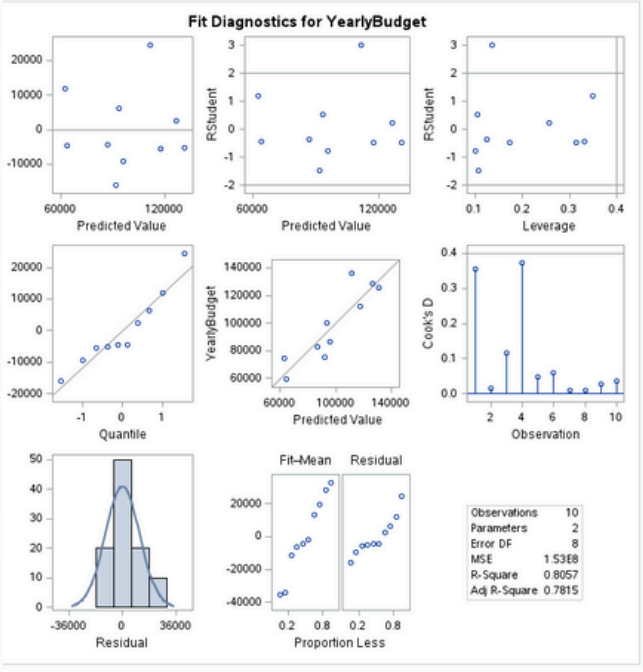
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	51699	8965.88808	5.77	0.0004
TotalEmployees	1	1.59884	0.27755	5.76	0.0004





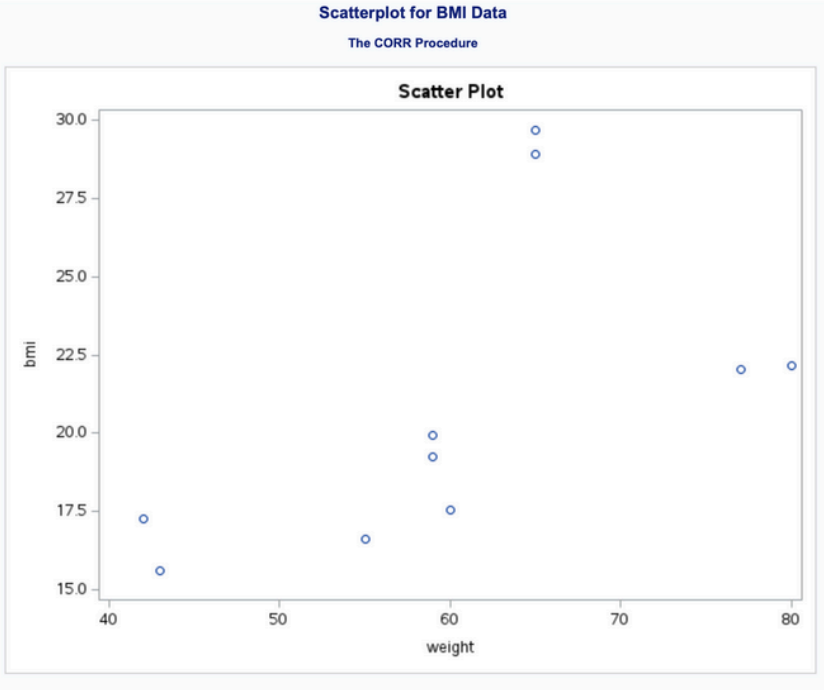
Straight-line Regression for Job Parametrics Data

The REG Procedure
Model: MODEL1
Dependent Variable: YearlyBudget



Code-Output

```
data data3;
 input gender $ age weight height;
 height2=height*height;
 bmi=weight/height2;
 datalines;
M 25 60 1.85
F 30 65 1.50
F 35 80 1.90
F 28 55 1.82
M 40 43 1.66
M 25 59 1.75
F 30 65 1.48
F 35 77 1.87
F 28 59 1.72
M 40 42 1.56
;
run;
ods graphics on;
ods select ScatterPlot;
proc corr data=data3 plots=scatter(noinset ellipse=none);
 var weight bmi;
 title 'Scatterplot for BMI Data';
run;
ods graphics off;
```



WORK.DATA3 | View: Column names | Filter: (none)

Total rows: 10 Total columns: 6 | Rows 1-10

gender	age	weight	height	height2	
M	25	60	1.85	3.4225	17.53
F	30	65	1.5	2.25	28.88
F	35	80	1.9	3.61	22.1
F	28	55	1.82	3.3124	16.60
M	40	43	1.66	2.7556	15.60
M	25	59	1.75	3.0625	19.26
F	30	65	1.48	2.1904	29.67
F	35	77	1.87	3.4969	22.01
F	28	59	1.72	2.9584	19.94
M	40	42	1.56	2.4336	17.25

Code-Output

```
proc reg data=data3;
 id bmi;
 model bmi=height height2 / p cli clm;
 title 'Fitting a Curve to the BMI Data';
 run;
 ods select OutputStatistics;
```

Fitting a Curve to the BMI Data

The REG Procedure
Model: MODEL1
Dependent Variable: bmi

Number of Observations Read	10
Number of Observations Used	10

Analysis of Variance

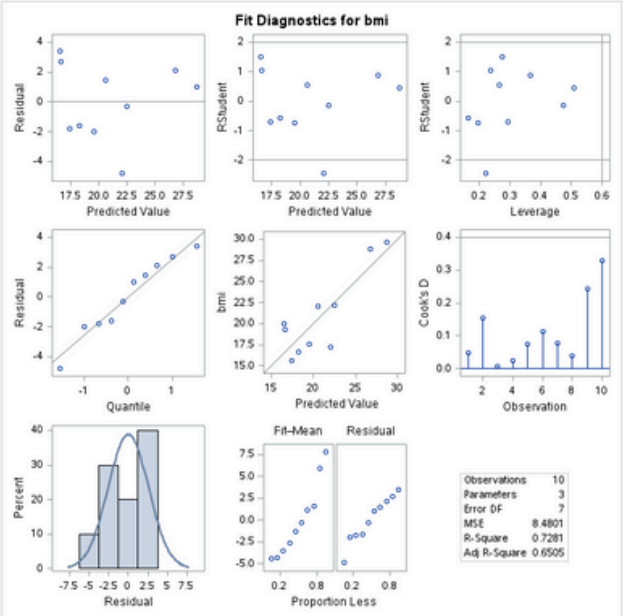
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	158.99702	79.49851	9.37	0.0105
Error	7	59.36086	8.48012		
Corrected Total	9	218.35789			

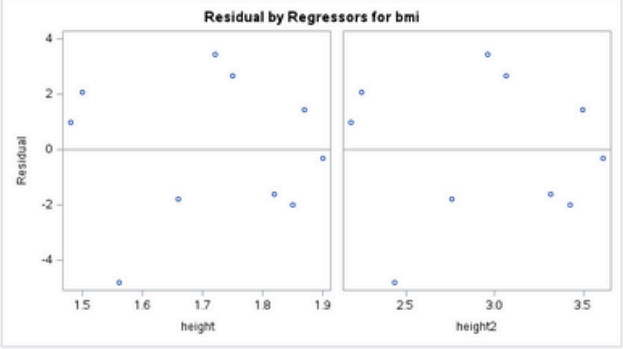
Root MSE	2.91207	R-Square	0.7281
Dependent Mean	20.89508	Adj R-Sq	0.6505
Coeff Var	13.93661		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	613.77995	162.55967	3.78	0.0069
height	1	-691.72695	193.93594	-3.57	0.0091
height2	1	200.27598	57.45469	3.49	0.0102

Sum of Residuals	-1.7518E-12
Sum of Squared Residuals	59.36086
Predicted Residual SS (PRESS)	108.04201





Fitting a Curve to the BMI Data

The REG Procedure
Model: MODEL1
Dependent Variable: bmi

Output Statistics

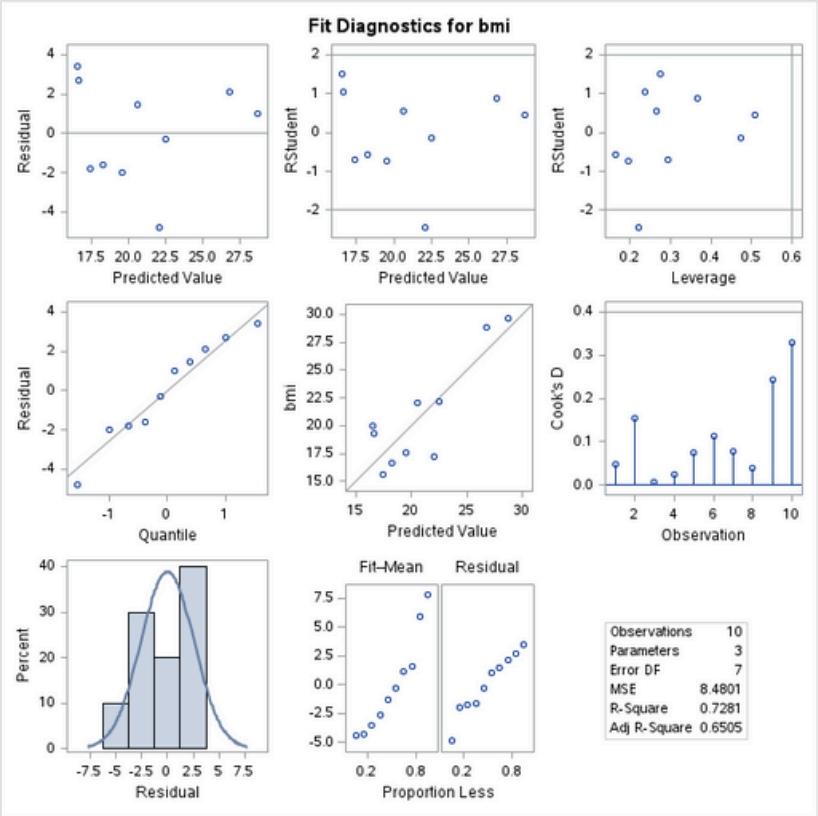
os	bmi	Dependent Variable	Predicted Value	Std Error Mean Predict	95% CL Mean		95% CL Predict		Residual
1	17.5	17.5	19.5296	1.2902	16.4789	22.5804	11.9982	27.0611	-1.9986
2	28.9	28.9	26.8105	1.7644	22.6384	30.9826	18.7593	34.8617	2.0784
3	22.2	22.2	22.4950	2.0027	17.7594	27.2307	14.1378	30.8523	-0.3344
4	16.6	16.6	18.2311	1.1751	15.4524	21.0098	10.8056	25.6565	-1.6268
5	15.6	15.6	17.3937	1.5766	13.6657	21.1217	9.5634	25.2241	-1.7891
6	19.3	19.3	16.6030	1.4190	13.2477	19.9583	8.9431	24.2629	2.6623
7	29.7	29.7	28.7086	2.0783	23.7943	33.6229	20.2489	37.1683	0.9664
8	22.0	22.0	20.5956	1.5011	17.0460	24.1453	12.8486	28.3426	1.4239
9	19.9	19.9	16.5061	1.5277	12.8936	20.1185	8.7301	24.2820	3.4371
10	17.3	17.3	22.0775	1.3659	18.8478	25.3073	14.4718	29.6833	-4.8192

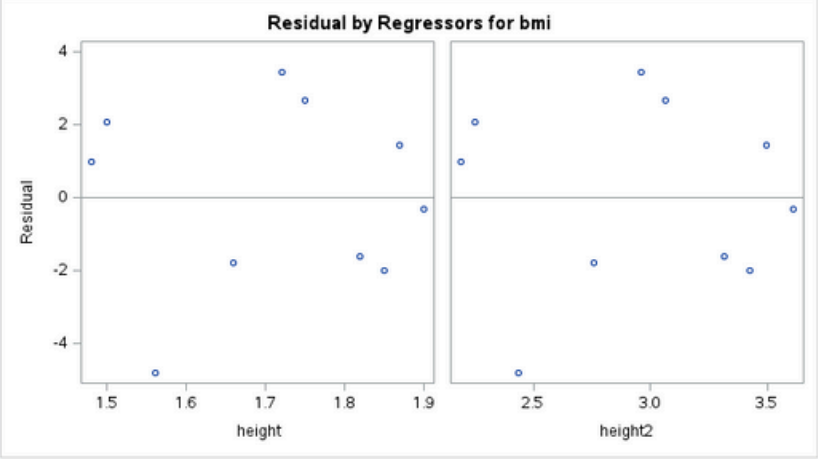
Sum of Residuals	-1.7518E-12
Sum of Squared Residuals	59.36086
Predicted Residual SS (PRESS)	108.04201

Code-Output

```
proc reg data=data3 alpha=0.1;
 model bmi=height height2 / p cli clm;
 title '90% Limits for BMI Data';
 run;
 ods graphics on;
 ods select PredictionPlot;
```

Sum of Residuals	-1.7518E-12
Sum of Squared Residuals	59.36086
Predicted Residual SS (PRESS)	108.04201





90% Limits for BMI Data

The REG Procedure
Model: MODEL1
Dependent Variable: bmi

Number of Observations Read	10
Number of Observations Used	10

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	158.99702	79.49851	9.37	0.0105
Error	7	59.36086	8.48012		
Corrected Total	9	218.35789			

Root MSE	2.91207	R-Square	0.7281
Dependent Mean	20.89508	Adj R-Sq	0.6505
Coeff Var	13.93661		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	613.77995	162.55967	3.78	0.0069
height	1	-691.72695	193.93594	-3.57	0.0091
height2	1	200.27598	57.45469	3.49	0.0102

90% Limits for BMI Data

The REG Procedure
Model: MODEL1
Dependent Variable: bmi

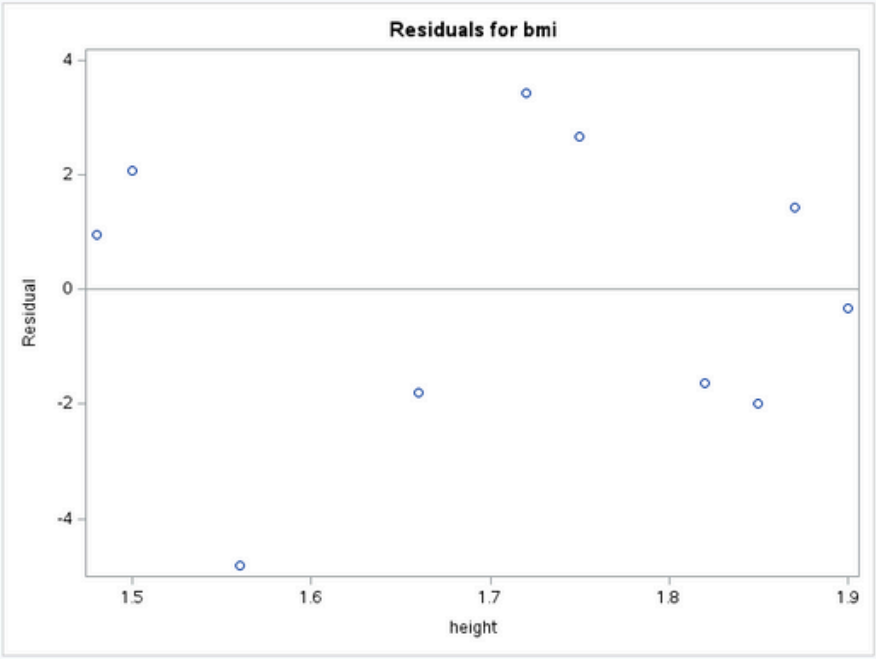
Output Statistics							
Dependent Variable	Predicted Value	Std Error Mean Predict	90% CL Mean		90% CL Predict		Residual
17.5	19.5296	1.2902	17.0854	21.9739	13.4953	25.5640	-1.9986
28.9	26.8105	1.7644	23.4677	30.1532	20.3597	33.2613	2.0784
22.2	22.4950	2.0027	18.7007	26.2894	15.7991	29.1910	-0.3344
16.6	18.2311	1.1751	16.0047	20.4574	12.2817	24.1805	-1.6268
15.6	17.3937	1.5766	14.4068	20.3807	11.1199	23.6675	-1.7891
19.3	16.6030	1.4190	13.9147	19.2913	10.4657	22.7403	2.6623
29.7	28.7086	2.0783	24.7712	32.6460	21.9305	35.4866	0.9664
22.0	20.5956	1.5011	17.7516	23.4397	14.3886	26.8027	1.4239
19.9	16.5061	1.5277	13.6117	19.4004	10.2758	22.7363	3.4371
17.3	22.0775	1.3659	19.4898	24.6653	15.9837	28.1714	-4.8192

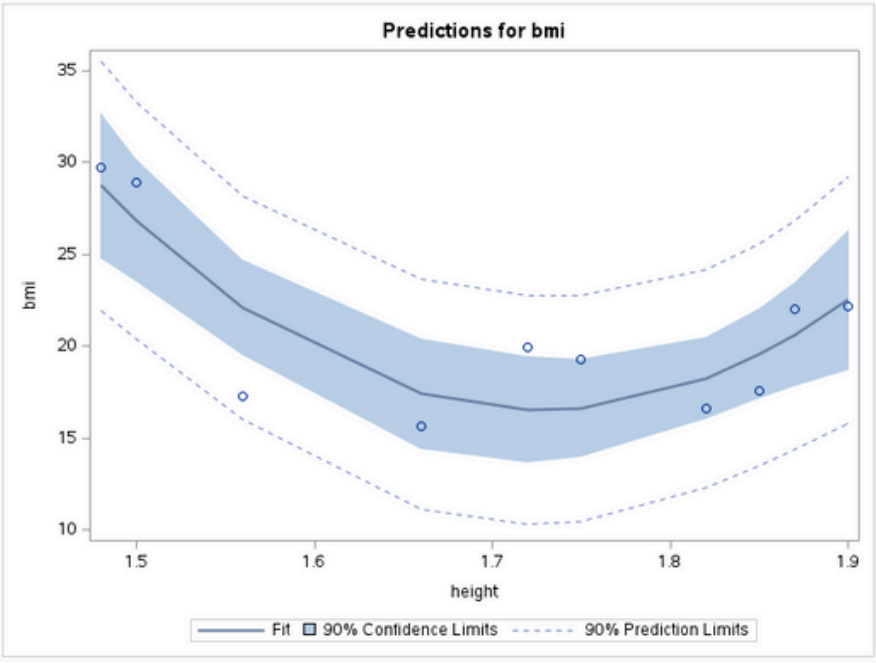
Sum of Residuals	-1.7518E-12
Sum of Squared Residuals	59.36086
Predicted Residual SS (PRESS)	108.04201

Code-Output

```
proc reg data=data3 alpha=0.10 plots(only)=predictions(x=height unpack);
 model bmi=height height2;
 run;
quit;
```

The REG Procedure
Model: MODEL1
Dependent Variable: bmi





The REG Procedure
Model: MODEL1
Dependent Variable: bmi

Number of Observations Read	10
Number of Observations Used	10

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	158.99702	79.49851	9.37	0.0105
Error	7	59.36086	8.48012		
Corrected Total	9	218.35789			

Root MSE	2.91207	R-Square	0.7281
Dependent Mean	20.89508	Adj R-Sq	0.6505
Coeff Var	13.93661		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	613.77995	162.55967	3.78	0.0069
height	1	-691.72695	193.93594	-3.57	0.0091
height2	1	200.27598	57.45469	3.49	0.0102