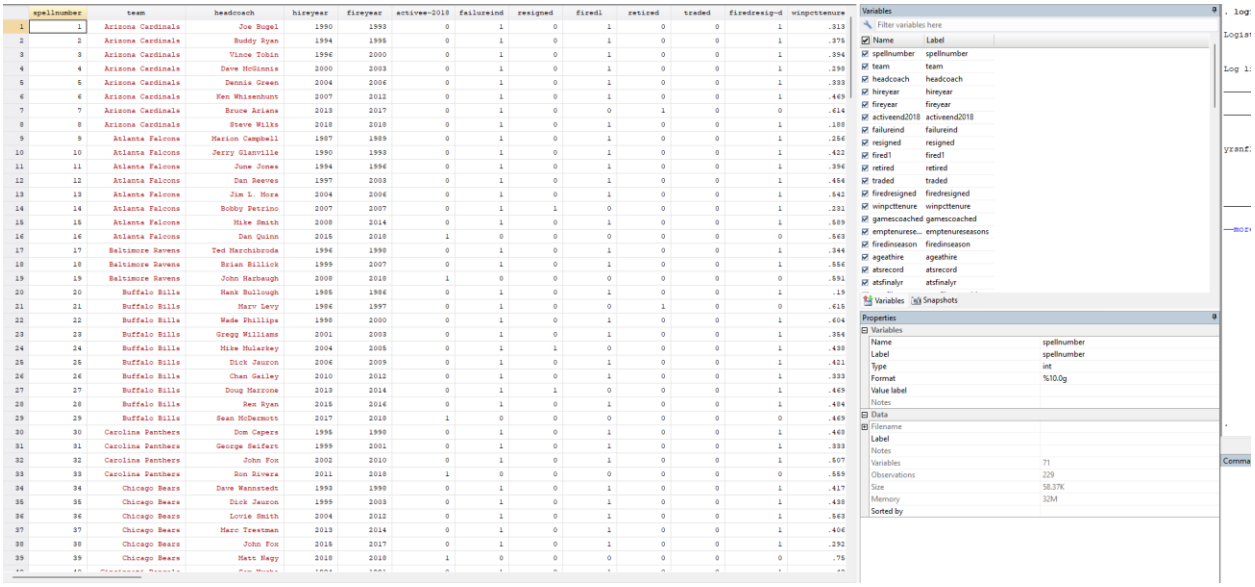
1-2)

I imported the data in Stata using the import excel command, I also needed to specify the sheet number and first row, so that Stata can read it as headers.



3-4)

I used summ command to get the summary of variables. Summaries provide us with the information of the number of observations, mean, standard deviation of the variable, minimum and maximum value of the variable. Summary of fired1, offcoach, and black is not relevant because they are discrete variables, and mean and standard deviations are not related but all other outputs can be useful in many cases.

```
. summ fired1 winpcttenure gamescoached ageathire yrsnflhcexpathire playedinnfl offcoach black
```

Variable	Obs	Mean	Std. Dev.	Min	Max
fired1	216	.7222222	.4489436	0	1
winpcttenure	216	.4402269	.1485205	.063	.766
gamescoached	216	67.56481	51.27469	13	304
ageathire	216	49.31019	7.185327	31	65
yrsnflhcex~e	216	2.377025	4.022251	0	17
playedinnfl	216	.3009259	.459726	0	1
offcoach	216	.6111111	.4886304	0	1
black	216	.1203704	.32615	0	1

We can identify the percentage of the white population by the frequency table of the white variable. I used tabulate command to create the frequency table. As we can see 86.57% of the population is white.

```
. tabulate white
```

white	Freq.	Percent	Cum.
0	29	13.43	13.43
1	187	86.57	100.00
Total	216	100.00	

.

5)

Since one is a continuous variable and the other is categorical, that's why I used the one-way ANOVA method.

```
. oneway yrsnflhcexpathire qb
```

Analysis of Variance

Source	SS	df	MS	F	Prob > F
Between groups	13.3112309	1	13.3112309	0.82	0.3656
Within groups	3465.06679	214	16.1919009		
Total	3478.37802	215	16.1785024		

```
Bartlett's test for equal variances:  chi2(1) =   1.4524  Prob>chi2 = 0.228
```

6)

Since we were willing to identify the determents of the fired variable and it's a binary variable, that's why I used a logistic regression model to model the relationship with independent variables. Also, I used odd ratios instead of coefficients because we are not interested in prediction but to interpret the result.

```
. logistic fired1 gmreplaced winpcttenure yrsnflhcexpathire careeravplayer coachgm laborpool

Logistic regression                             Number of obs     =        144
                                                LR chi2(6)        =      41.58
                                                Prob > chi2       =     0.0000
Log likelihood = -59.069667                     Pseudo R2         =     0.2603
```

fired1	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
gmreplaced	2.629664	1.157188	2.20	0.028	1.110012	6.229783
winpcttenure	.0000654	.0001476	-4.27	0.000	7.86e-07	.005443
yrsnflhcexpathire	1.015552	.0541158	0.29	0.772	.9148377	1.127353
careeravplayer	1.012287	.0132001	0.94	0.349	.9867432	1.038492
coachgm	.4340821	.2652649	-1.37	0.172	.1310412	1.437924
laborpool	1.425645	.1932698	2.62	0.009	1.092992	1.859542
_cons	.983368	2.198738	-0.01	0.994	.0122883	78.69367

I selected the following variables for determinants.

Variable	Reason
gmreplacted	Gmreplaced provides the information about the replacement of the manager during the head coach employment spell. GM of the team is replaced when the club decides on a fresh start, and they might be interested in firing the coach as well.
Winpcttenure	Winpcttenure provides information about the winning percentage of the team during employment spell. Since winning percentage is the main metric that decides the performance of the team, that's why it's directly related to the promotions and demotions/fired of the head coach.
Yrsnflhcexpathire	The experience of the head coach is one of the main factors related to the performance of the team. The performance of the team is related to the firing of the head coach. That's why I selected this variable as well.
Careeravplayer	A good player can become a better coach in the future. That's why it's necessary to keep the performance of the coach as a team player. That's why I included careeravplayer that presents the head coach career "Approximate Value" metric as a player.
Coachgm	Coaching a full team is one of the tough jobs but if someone is being a coach and general manager at the same time. It might affect the performance of the team and ends up firing the head coach.
laborpool	If there are a smaller number of applicants for the head coach it will defiantly decrease the quality of the head coach. That's why laborpool is the best variable to add because it provides information about the number of recent HC available for hire at the time HC lost his HC position

7)

According to the model statistics, 144 observations are included in the model. While the raw data has 217 observations. So, the number of observations is not same in the model and excel sheet.

Stata uses a listwise deletion approach in regression analysis that drops all observations that have a missing value for any one of the variables used in the model that's why our variable contains missing values, and those rows are deleted before model implementation.

8)

Gmreplacted:

It has a p-value of 0.028 which represents that it's statistically significant at 0.05 level. Its coefficient is 2.629664 which is also greater than 1 and shows a positive relation with fired1.

Winpcttenure:

It has a p-value of < 0.000 which represents that it's statistically significant at 0.05 level. Its coefficient is 0.0000654 which is less than 1 and shows a negative relation with fired1.

Yrsnflhcexpathire:

It has a p-value of 0.772 which represents that it's not statistically significant at 0.05 level. Its coefficient is 1.015552 which is almost one and shows no relation with fired1.

Careeravplayer:

It has a p-value of 0.349 which represents that it's not statistically significant at 0.05 level. Its coefficient is 1.015552 which is almost one and shows no relation with fired1.

Coachgm:

It has a p-value of 0.172 which represents that it's not statistically significant at 0.05 level. Its coefficient is 0.4340821 which is less than 1 and shows a negative relation with fired1.

Laborpool:

It has a p-value of 0.009 which represents that it's statistically significant at 0.05 level. Its coefficient is 1.425645 which is also positive and shows a positive relation with fired1.

9)

If I can add two more variables, I will add the college and draftednfl. Better college and relevant experience can help them to create a well-educated and well-experienced head coach which will help to increase the team performance.

1-2)

There's a variable winpcttenure which provides information about winning percentage over employment spell of the head coach. We can use multiple linear regression to model the relation of determinants with winpcttenure, which will help us to identify the determinants for team performance.

I used multiple linear regression to model the relation between winpcttenure (HC winning percentage over employment spell) and 6 independent variables. Independent variables are atsrecord, normpayspell, coachgm, teamwpcthire5, gmreplaced, and allexperience.

```
. regress winpcttenure atsrecord normpayspell coachgm teamwpcthire5 gmreplaced allexperience
```

Source	SS	df	MS		
				Number of obs	= 140
				F(6, 133)	= 28.85
Model	1.6250325	6	.27083875	Prob > F	= 0.0000
Residual	1.24867004	133	.009388497	R-squared	= 0.5655
				Adj R-squared	= 0.5459
Total	2.87370254	139	.020674119	Root MSE	= .09689

winpcttenure	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
atsrecord	1.513527	.1482953	10.21	0.000	1.220205	1.80685
normpayspell	.0495531	.013292	3.73	0.000	.0232621	.0758441
coachgm	.0252151	.0244461	1.03	0.304	-.0231383	.0735685
teamwpcthire5	.3083489	.0729171	4.23	0.000	.1641217	.4525762
gmreplaced	.0260781	.0149039	1.75	0.082	-.0034014	.0555575
allexperience	.00042	.001128	0.37	0.710	-.0018112	.0026512
_cons	-.4396223	.0841737	-5.22	0.000	-.6061147	-.27313

Atsrecod:

Head coach regular-season record can be helpful to identify the performance of the head coach and it can directly affect the team winning percentage.

Normpayspell:

Better payroll attracts good player to the team which increases the competition between players and increase the overall game performance for the team.

Coachgm:

HC and Gm are the core of any team and a good understanding between HC and GM always results in a good performance. But if the head coach and Gm are the same person, it might give a boost to the overall performance.

Teampcthire5:

Team regular-season winning percentage before hiring the HC and after HC can help us to identify if hiring the head coach help to increase the winning percentage or improve the performance of the team.

Gmreplaced:

Information about GM being replaced during the HC employment spell can help us to identify that can changing the GM increase the winning percentage or decrease the winning percentage?

Allexperience:

The coaching experience of the head coach may have an impact on the performance of the team.

3)

According to the model statistics, 140 observations are included in the model. While the raw data has 217 observations. So, the number of observations in the model in the excel sheet is not the same.

Stata uses a listwise deletion approach in regression analysis that drops all observations that have a missing value for any one of the variables used in the model that's why our variable contains missing values, and those rows are deleted before model implementation.

4)

```
. regress winpcttenure atsrecord normpayspell coachgm teamwpcthire5 gmreplaced allexperience
```

Source	SS	df	MS	Number of obs	=	140
				F(6, 133)	=	28.85
Model	1.6250325	6	.27083875	Prob > F	=	0.0000
Residual	1.24867004	133	.009388497	R-squared	=	0.5655
				Adj R-squared	=	0.5459
Total	2.87370254	139	.020674119	Root MSE	=	.09689

winpcttenure	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
atsrecord	1.513527	.1482953	10.21	0.000	1.220205	1.80685
normpayspell	.0495531	.013292	3.73	0.000	.0232621	.0758441
coachgm	.0252151	.0244461	1.03	0.304	-.0231383	.0735685
teamwpcthire5	.3083489	.0729171	4.23	0.000	.1641217	.4525762
gmreplaced	.0260781	.0149039	1.75	0.082	-.0034014	.0555575
allexperience	.00042	.001128	0.37	0.710	-.0018112	.0026512
_cons	-.4396223	.0841737	-5.22	0.000	-.6061147	-.27313

Atsrecord:

atsrecord has a p-value < 0.001 which represents that atsrecord is statistically significant at 0.05 level. The coefficient of ats record is 1.513527. Which shows a positive relation between atsrecord and winpcttenure.

Normpayspell:

normpayspell has p-value < 0.001 which represents that it is statistically significant at 0.05 level. The coefficient of normpayspell is 0.0495531 which is also positive and shows a positive relation normpayspell and winpcttenure.

Teamwpcthire5:

It has a p-value of < 0.001 which represents that it's statistically significant at 0.05 level. Its coefficient is 0.3083489 which is also positive and shows a positive relation with winpcttenure.

Coachgm:

It has a p-value of 0.304 which represents that it's not statistically significant at 0.05 level. Its coefficient is 0.0252151 which is also positive and shows a positive relation between coachgm and winpcttenure.

Gmreplaced:

It has a p-value of 0.082 which represents that it's not statistically significant at 0.05 level. Its coefficient is 0.0260781 which is also positive and shows a positive relation with winpcttenure.

Allexperience:

It has a p-value of 0.710 which represents that it's not statistically significant at 0.05 level. Its coefficient is 0.00042 which is also positive and shows a positive relation with winpcttenure.

5)

If I can add two more variables, I will add the college and yearnflvariable. Better college and relevant experience can help them to create a well-educated and well-experienced head coach which will help to increase the team performance.