Assignment 3

PAI 721: Introduction to Statistics

Prof. Jack Reilly

F2025

Instructions

For this assignment, you must turn in two documents:
1. Your answers (as a PDF), and
2. A plain-text `.do` file containing the Stata code you used to arrive at your answers.

Please consider the data in the following table. You may also find this data as a `.dta` file on the course Blackboard page.

Dataset

Player	G	MP	FG	FGA	3P	3PA	TRB	AST	STL	BLK	TOV	PF	PTS
Carmelo Anthony	35	1274	277	612	56	166	349	77	55	30	77	77	778
Hakim Warrick	35	1146	197	364	0	1	297	57	49	44	92	94	518
Gerry McNamara	35	1236	146	364	85	238	80	155	77	2	85	69	467
Kueth Duany	35	944	133	303	43	123	128	71	36	17	57	75	386
Billy Edelin	23	533	80	146	0	2	78	58	24	2	53	22	208
Josh Pace	32	469	62	118	0	2	86	60	26	8	37	33	138
Craig Forth	35	618	56	115	0	1	116	30	15	41	39	90	132
Jeremy McNeil	35	657	54	81	0	0	146	8	9	100	36	112	117
Matt Gorman	9	76	8	23	0	1	19	1	3	2	5	14	21

Player	G	MP	FG	FGA	3P	3PA	TRB	AST	STL	BLK	TOV	PF	PTS
Andrew Kouwe	6	10	3	5	2	2	2	2	0	0	1	2	10
Ronneil Herron	5	9	2	3	0	0	5	0	0	0	1	0	6
Xzavier Gaines	6	21	1	8	0	3	2	1	1	0	2	0	2
Gary Hall	5	6	1	1	0	0	2	2	1	1	0	0	2
Tyrone Albright	7	20	0	2	0	1	2	1	2	0	3	1	0
Josh Brooks	5	6	0	1	0	0	1	0	0	0	0	1	0

Variable definitions:

- `G` = Games Played
- `MP` = Minutes Played
- `FG` = Field Goals Made
- `FGA` = Field Goals Attempted
- `3P` = Three Point Field Goals Made
- `3PA` = Three Point Field Goals Attempted
- `TRB` = Total Rebounds
- `AST` = Assists
- `STL` = Steals
- `BLK` = Blocks
- `TOV` = Turnovers
- `PF` = Personal Fouls
- `PTS` = Points Scored

Questions

1. What is the unit of analysis (the cases) in this table?
2. Identify the level of measurement of each variable.
3. What is the mean, median, and mode of the "games played" variable in this data?
4. Do you think the mean, median, or mode is the best estimate of central tendency for games played? Why?
5. Many analysts like to create a statistic called "Stocks," which is the sum of steals and blocks by a player. Create this variable as `STK`. Who had the most stocks on the team?
6. Calculate each player's field goal percentage and three point field goal percentage, storing each in new variables `FGP` and `TPP`. Who led the team in field goal percentage and three point percentage?
7. Transform the variable `PTS`, generating a new variable `PPG` (average points scored per game played). How many points per game did Carmelo Anthony, Gerry McNamara, Josh Pace, and Jeremy McNeil average? Who led the team in points per game?
8. Do the same for `3P`, `TRB`, `AST`, `STL`, `BLK`, and `TOV`. Who led the Orange in each statistical category per game?
9. Transform `PTS` again to generate `PPM` (points per minute played). Who led the Orange in points per minute?
10. Create a scatterplot with points per minute (`PPM`) on the Y axis and minutes played (`MP`) on the X axis. Add a best-fit line. What is the relationship between points per minute and minutes played?