

AP Statistics Exploratory Data Analysis

1. **Cholesterol** A study examining the health risks of smoking measured the cholesterol levels of people who had smoked for at least 25 years and people of similar ages who had smoked for no more than 5 years and then stopped. Create a graphical display for both groups and write a report comparing their cholesterol levels.

Smokers				Ex-smokers		
225	2111	209	284	250	134	300
258	216	196	288	249	213	310
250	200	209	280	175	174	328
225	256	243	200	160	188	321
213	246	225	237	213	257	292
232	267	232	216	200	271	227
216	243	200	155	238	163	263
216	271	230	309	192	242	249
183	280	217	305	242	267	243
287	217	246	351	217	267	218
200	280	209		217	183	228

2. **Baseball** American League baseball teams play their games with a designated hitter rule, meaning that pitchers do not bat. The League believes that replacing the pitcher, typically a weak hitter, with another player in the batting order produces more runs and generates more interest among fans. Below are the average numbers of runs scored in American League and National League stadiums for the first half of the 2001 season.

American	11.1	10.8	10.8	10.3	10.3	10.1	10	9.5	9.4	9.3	9.2	9.2	9	8.3		
National	14	11.6	10.4	10.3	10.2	9.5	9.5	9.5	9.5	9.1	8.8	8.4	8.3	8.2	8.1	7.9

- Create a back-to-back stemplot of these data.
- Write a few sentences comparing the average number of runs scored per game in the two leagues.
- Coors Field, in Denver, stands a mile above sea level, an altitude far greater than that of any other major league ball park. Some believe that the thinner air makes it harder for pitchers to throw curve balls and easier for batters to hit the ball a long way. Do you see any evidence that the 14 runs scored per game there is unusually high? Explain.

3. **Drunk Driving** Accidents involving drunk drivers account for about 40% of all deaths on the nation's highways. The table tracks the number of alcohol-related fatalities for 20 years.

Year	Death (thousands)	Year	Death (thousands)
1982	25.2	1992	17.9
1983	23.6	1993	17.5
1984	23.8	1994	16.6
1985	22.7	1995	17.2
1986	24	1996	17.2
1987	23.6	1997	16.5
1988	23.6	1998	16
1989	22.4	1999	16
1990	22	2000	16.7
1991	19.9	2001	16.7

- Create a stemplot or a histogram of these data.
- Create a timeplot.
- Using features apparent in the stemplot (or histogram) and the timeplot, write a few sentences about deaths caused by drunk driving.