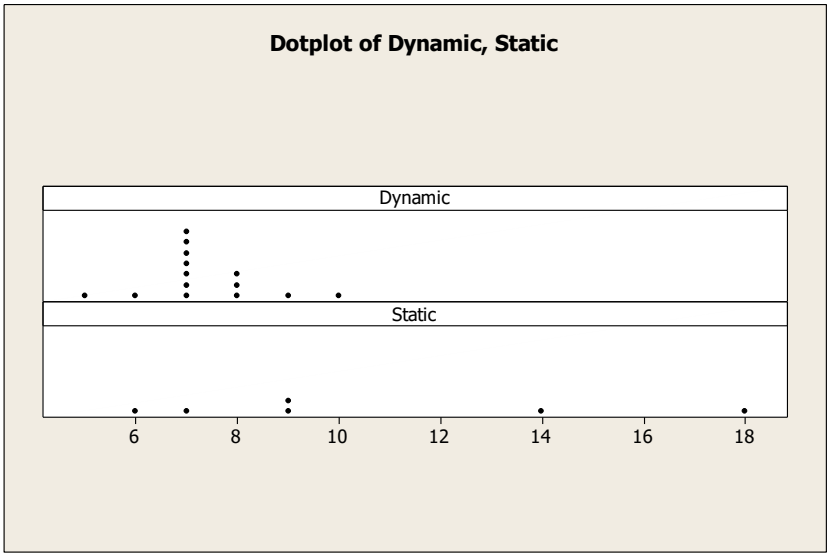
MAT162 Introduction to Applied Statistics

Description: Graphical and quantitative description of data; binomial, normal and t-distributions; one and two sample hypothesis tests and confidence intervals; simple linear regression and correlation. *Prerequisite (C or better):* MAT097 or MAT121 or higher. Prerequisite or corequisite: RDG100

Learning Outcomes	Sample Problems
1. (Comprehension Level) Explain simple statistical methods commonly used in reporting polling data and scientific research studies using correct statistical notation and appropriate language.	Discuss the sampling methods used by Gallop and the Literary Digest during the election of 1936.
2. (Synthesis Level) Construct informative graphical and numerical summaries of data appropriate for the type of data and the context in which the data was collected.	Age of onset of diabetes (35 people): 48 41 57 83 41 55 59 61 38 48 79 75 77 7 54 23 47 56 79 68 61 64 45 53 82 68 38 70 10 60 83 76 21 65 47 *What type of graph is appropriate for this data?* *What story is the graph telling us?*
3. (Evaluation Level) Interpret the meaning of graphical and numerical summaries of data in written terms appropriate to the context in which the data was collected.	Several neurosurgeons wanted to determine whether a dynamic system reduced the number of postoperative days in the hospital relative to the static system. Use the dotplots below to compare the two data sets. 
4. (Analysis Level) Recognize and properly carry out parameter estimation and hypothesis testing	The average starting salary for a radiographer is $44,485. A radiology program at a local community college believes its graduates begin at higher salaries. They randomly sample 12 recent graduates and find their

<table>
<tr><td>procedures with and without the use of technology.</td><td>average starting salary was $47,000 with a standard deviation of $7,800. Assume the starting salary for graduates of this program is approximately normal. At the 5% significance level, does the community college's claim have any merit? Use the critical value approach and show all of your work. Be sure to clearly state your conclusion using complete sentences.</td></tr>
<tr><td>5. (Comprehension Level) Discuss the formalism of parameter estimation and hypothesis testing and how it relates to, supports and advances the scientific method.</td><td>Explain the general procedure for deciding whether a null hypothesis should be rejected? Explain why we can reject a null hypothesis, but we can never accept or prove it? Finally, discuss why statistics is known as the "language of science".</td></tr>
<tr><td>6. (Analysis Level) Apply parameter estimation and hypothesis testing to solve problems utilizing appropriate statistical methods.</td><td>A biologist was interested in determining whether specific seedlings treated with a new fertilizer resulted in a higher than average height than the standard height of 11.2 cm. The biologist treated a random sample of 24 seedlings with the fertilizer and subsequently obtained the following heights. Conduct an appropriate hypothesis test using a 5% significance level.

<table><tr><td>11.5</td><td>11.8</td><td>15.7</td><td>12.1</td><td>10.1</td><td>10.5</td></tr><tr><td>16.5</td><td>13.5</td><td>14.4</td><td>16.7</td><td>10.9</td><td>10.0</td></tr><tr><td>11.1</td><td>17.1</td><td>13.3</td><td>12.4</td><td>8.5</td><td>14.3</td></tr><tr><td>12.9</td><td>11.1</td><td>15.0</td><td>13.3</td><td>9.9</td><td>13.5</td></tr></table></td></tr>
<tr><td>7. (Analysis Level) Recognize the limitations of statistical methods and discuss the appropriateness of use within a context.</td><td>A clean air standard requires that vehicle exhaust emissions not exceed specified limits for various pollutants. Many states require cars be tested annually to be sure they meet these standards. Suppose state regulators double check a random sample of cars that a suspect repair shop has certified as okay. They will revoke the shop's license if they find significant evidence that the shop is certifying vehicles that do not meet standards. a) What are the null and alternative hypotheses? b) In this context, what is a Type I error? c) In this context, what is a Type II error? d) Which type of error would the shop's owner consider more serious? Why? e) Which type of error might environmentalists consider more serious? Why?</td></tr>
<tr><td>8. (Application Level) Empirically and theoretically obtain the probability of an event.</td><td>A certain rare blood type is found in only 1% of the population. If 300 people from the population are randomly selected, what is the probability that at least one of them will have the rare blood type?</td></tr>
<tr><td>9. (Application Level) Apply the normal distribution to calculate probability of event.</td><td>Suppose that the incubation period—that is, the time between being infected with the virus and showing symptoms—for the Ebola virus is normally distributed with mean $\mu = 12$ days and a standard</td></tr>
</table>

	deviation of σ = 3 days. Find the probability that a randomly selected case demonstrated symptoms in fewer than 7 days.
10. (Application Level) Apply the properties of the t-distribution, t-statistic, and degrees of freedom to construct confidence intervals.	A hospital wants to estimate the average length of post-op stay. The *standard deviation of the lengths of stay of all previous surgery patients is known to be 4.1 days*. A random sample of 50 surgery patients' records is obtained, and the mean length of time these patients stayed in the hospital was 6.2 days. Find a 98% confidence interval for the mean length of time all surgery patients stay in the hospital.
11. Examine the relationship between two quantitative variables using the correlation coefficient and by computing the regression equation.	Use the data below to obtain the correlation coefficient and interpret r. Age (x): 10 12 12 14 18 18 20 23 26 27 Length (y): 64 67 110 104 162 166 177 228 231 279