

Chemical Marketing Report's 2016 list of 100 leading companies in petrochemicals, ExxonMobil is the largest company in the world by revenue, profit, and market value. The company's total revenue in 2016 was \$362.5 billion, up 1.6% from 2015. Total assets were \$304.9 billion, up 1.6% from 2015. Total net income was \$26.8 billion, up 1.6% from 2015. Total cash flow was \$33.2 billion, up 1.6% from 2015. Total earnings per share were \$4.00, up 1.6% from 2015. Total dividends per share were \$1.60, up 1.6% from 2015. Total market value was \$275.4 billion, up 1.6% from 2015. Total debt was \$100.5 billion, up 1.6% from 2015.

## SECTION II PART A

### Questions 1–5

#### Question One

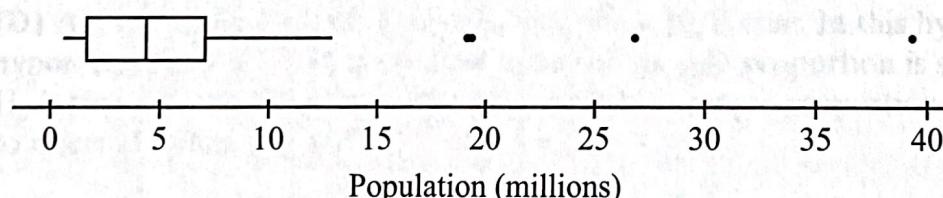
(a)  $IQR = 7.0 - 1.6 = 5.4$ .

$$Q_1 - 1.5(IQR) = 1.6 - 1.5(5.4) = -6.5$$

$$Q_3 + 1.5(IQR) = 7.0 + 1.5(5.4) = 15.1$$

There are no populations less than  $-6.5$ . Any population greater than  $15.1$  is an outlier, so the outliers are  $19.7$ ,  $19.9$ ,  $27.0$ , and  $38.8$ .

(b)



- (c) Missouri's population is greater than the median population, so Missouri is among the largest fifty percent of states. However, its population is less than the upper quartile population, so Missouri is not in the largest 25% of states.

**Question Two**

(a)

|                    |   | Score on Spinner 1 |   |   |   |   |   |
|--------------------|---|--------------------|---|---|---|---|---|
|                    |   | 1                  | 2 | 3 | 4 | 5 | 6 |
| Score on Spinner 2 | 1 | 0                  | 1 | 2 | 3 | 4 | 5 |
|                    | 2 | 1                  | 0 | 1 | 2 | 3 | 4 |
|                    | 3 | 2                  | 1 | 0 | 1 | 2 | 3 |
|                    | 4 | 3                  | 2 | 1 | 0 | 1 | 2 |

(b)

| Amount Won (\$) | 0   | 1    | 2   | 3   | 4    | 5    |
|-----------------|-----|------|-----|-----|------|------|
| Probability     | 1/6 | 7/24 | 1/4 | 1/6 | 1/12 | 1/24 |

(c)  $E(\text{Amount won}) = 0(1/6) + 1(7/24) + 2(1/4) + 3(1/6) + 4(1/12) + 5(1/24) = 1.83$ .

(d) The stallholder charges \$2.50 for each play and, on average, the participant wins \$1.83. So, on average, the stallholder makes  $2.50 - 1.83 = \$0.67$  per play of the game. So, in 400 plays, the stallholder can expect to make  $400(0.67) = \$268$ .



### **Question Three**

- (a) She could number the hospitals 1–439, and then use the computer's random number generator to select 40 different numbers in that range. The 40 hospitals with these numbers would form the sample.
  - (b) Randomly select 10 urban teaching hospitals, 10 urban non-teaching hospitals, 10 rural teaching hospitals, and 10 rural non-teaching hospitals. These 40 hospitals form the sample.
  - (c) It could be suggested that urban hospitals are more expensive than rural hospitals and teaching hospitals are more expensive than non-teaching hospitals.

Unaú

thoriz

ed co

copyin

10

euusim

6

/ part

of thi

100

Page 11

**Question Four**

| Student | No Music | Music | Difference<br>(No Music – Music) |
|---------|----------|-------|----------------------------------|
| 1       | 26       | 27    | 1                                |
| 2       | 10       | 7     | -3                               |
| 3       | 19       | 20    | 1                                |
| 4       | 25       | 28    | 3                                |
| 5       | 8        | 5     | -3                               |
| 6       | 23       | 24    | 1                                |
| 7       | 21       | 24    | 3                                |
| 8       | 13       | 12    | -1                               |
| 9       | 9        | 11    | 2                                |
| 10      | 26       | 26    | 0                                |
| 11      | 10       | 12    | 2                                |
| 12      | 25       | 24    | -1                               |

Hypotheses:

Let  $\mu_d$  be the true mean difference in on-task time (music minus no music).

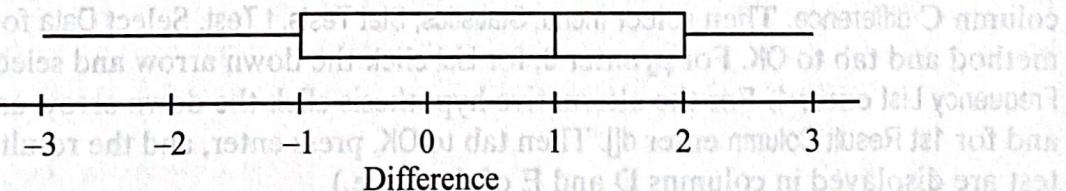
$$H_0: \mu_d = 0$$

$$H_a: \mu_d > 0$$

Check of conditions and naming of test:

This is a paired  $t$ -test.

We are told that the students were randomly assigned to the two orders of treatments. Since this is a small sample we consider a boxplot of the differences:





There are no outliers in the differences and the boxplot is close enough to being symmetrical for a sample this small. So we can proceed with the paired  $t$ -test.

### Mechanics:

$$t = \frac{\bar{d} - \mu_d}{\left( \frac{s_d}{\sqrt{n}} \right)} = \frac{0.417 - 0}{\left( \frac{2.065}{\sqrt{12}} \right)} = 0.699.$$

We use the  $t$ -distribution with  $12 - 1 = 11$  degrees of freedom. So the  $p$ -value is  $P(t_{11} > 0.699) = 0.250$ .

(This  $p$ -value can be obtained on the TI-83, TI-84, and TI-Nspire calculators by entering `tcdf(0.699,999,11)`. The `tcdf` function is accessed on the TI-83/84 by 2nd,DISTR, and on the TI-Nspire, in the *Scratchpad* or *Calculator* application, `tCdf` is accessed by selecting menu, Statistics, Distributions.)

### Conclusion:

Since the  $p$ -value is 0.250, which is greater than 0.05, we do not reject  $H_0$ . The results do not provide convincing evidence that the true mean on-task time when music is playing is greater than when there is no music.

(The numerical results of this hypothesis test can be found directly, using a calculator, as follows. On the TI-83 and TI-84 calculators, first select STAT, Edit and enter the scores for last year into L1 and the scores for this year into L2. Go to the home screen and type `L2-L1→L3`; the differences are then stored in L3. Then select STAT, TESTS, T-Test. For Inpt select Data, for  $\mu_0$  enter 0, for List enter L3, and for Freq enter 1. Select  $>\mu_0$  for the alternative hypothesis, select Calculate, and press Enter. The results of the test are now displayed.)

On the TI-Nspire, from the home screen tab to the Add Lists & Spreadsheet icon. Press enter, and enter the two data sets into columns A and B. Name the lists `list1` and `list2`, respectively. Move to the formula cell for column C, press enter, and then type `list2-list1`, and press enter. Name column C difference. Then select menu, Statistics, Stat Tests, t Test. Select Data for the data input method and tab to OK. For  $\mu_0$  enter 0, for List click the down arrow and select difference. For Frequency List enter 1. For the alternative hypothesis click the down arrow and select  $H_a: \mu > \mu_0$ , and for 1st Result Column enter `d1`. Then tab to OK, press enter, and the results of the hypothesis test are displayed in columns D and E of the table.)

**Question Five**

- (a) The equation of the least squares regression line is  $\hat{y} = 7.418 - 0.01670x$ , where  $x$  = temperature and  $y$  = pH.
- (b) The predicted pH when  $x = 30$  is  $\hat{y} = 7.418 - 0.01670(30) = 6.917$ . So the residual for the given observation is  $7.08 - 6.917 = 0.163$ . The pH for this observation was 0.163 higher than the predicted pH for water at  $30^\circ\text{C}$ .
- (c) Yes. The  $p$ -value for the  $t$ -test for the slope of the population regression line is given by the computer output to be 0.007. The computer output always supplies the  $p$ -value for a two-tailed test, and the fact that this is less than 0.05 (or any commonly-used significance level) provides us with convincing evidence of a linear relationship between temperature and pH.

**SECTION II PART B****Section A****Section B**

- (a) The Central Limit Theorem tells us that the sampling distribution of the sample mean  $\bar{X}$  is approximately normal.

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{1 + 0 + 3 + 4 + 2 + 1 + 0 + 1 + 3 + 2}{\sqrt{10}} = \frac{14}{\sqrt{10}} = 4.58$$

The Central Limit Theorem tells us that the sampling distribution of the sample mean  $\bar{X}$  is approximately normal. This means that the probability of obtaining a sample mean  $\bar{X}$  that is greater than or equal to 10.30 is approximately 0.10. This is because  $P(\bar{X} \geq 10.30) = P\left(\frac{\bar{X} - 10.0}{4.58} \geq \frac{10.30 - 10.0}{4.58}\right) = P(Z \geq 0.66)$ .

$$P(Z \geq 0.66) = 1 - P(Z \leq 0.66) = 1 - 0.7475 = 0.2525$$

$$P(Z \geq 0.66) = 1 - P(Z \leq 0.66) = 1 - 0.7475 = 0.2525$$

The Central Limit Theorem tells us that the sampling distribution of the sample mean  $\bar{X}$  is approximately normal. This means that the probability of obtaining a sample mean  $\bar{X}$  that is greater than or equal to 10.30 is approximately 0.10. This is because  $P(\bar{X} \geq 10.30) = P\left(\frac{\bar{X} - 10.0}{4.58} \geq \frac{10.30 - 10.0}{4.58}\right) = P(Z \geq 0.66)$ .

the sample mean is 3.15, which is close enough to being 3.5 to be considered a reasonable estimate.

The standard deviation of the sampling distribution of  $\bar{X}$  is given by the formula  $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ , where  $\sigma$  is the population standard deviation and  $n$  is the sample size. Here,  $\sigma = 1.708$  and  $n = 10$ , so  $\sigma_{\bar{X}} = \frac{1.708}{\sqrt{10}} = 0.540$ . The probability that the sample mean is less than or equal to 3.15 is given by the formula  $P(\bar{X} \leq 3.15) = P\left(Z \leq \frac{3.15 - 3.5}{0.540}\right) = P(Z \leq -0.6481)$ . This probability can be found using the TI-83/84 calculator by entering `normalcdf(-999,-0.6481,0,1)` or on the TI-Nspire by entering `normalCdf(-999,-0.6481,0,1)`.

## SECTION II PART B

### Question 6

#### Question Six

- The Central Limit Theorem tells us that if  $n$  is large the sampling distribution of  $\bar{X}$  is approximately normal.
- $\bar{X} = \frac{5 + 3 + 4 + 6 + 1 + 3 + 4 + 3 + 6 + 1}{10} = 3.6$
- The standard deviation of the sampling distribution of  $\bar{X}$  is  $\sigma/\sqrt{n}$ , where  $\sigma$  is the population standard deviation and  $n$  is the sample size. (Here, the population is the set of all possible scores when the die is rolled.) The population standard deviation is given to be 1.708, and  $n = 10$ . So the standard deviation of the sampling distribution of  $\bar{X}$  is  $1.708/\sqrt{10} = 0.540$ .
- $P(\bar{X} \leq 3.15) = P\left(Z \leq \frac{3.15 - 3.5}{0.540}\right) = P(Z \leq -0.6481) = 0.2584$ . So the number of values of  $\bar{X}$  that you would expect to be less than or equal to 3.15 is  $0.2584(100) = 25.84$ .

(The probability in the calculation above can be found using the TI-83 and TI-84 calculators by entering `normalcdf(-999,-0.6481)` or `normalcdf(-999,-0.6481,0,1)` and on the TI-Nspire by entering `normalCdf(-999,-0.6481,0,1)`. The `normalcdf` function is accessed on the TI-83/84 by 2nd DISTR, and the `normalCdf` function is accessed on the TI-Nspire, in the *Scratchpad* or *Calculator* application, by selecting menu, Statistics, Distributions. Alternatively, the answer can be found directly by entering `normalcdf(-999,3.15,3.5,0.540)`.)



(e) 95% of the state's CDs will be included in the sample.

- The expected count calculated in part (d) has been written in the table in part (ii).
- $100 - (25.84 + 21.52 + 16.75 + 15.42) = 20.47$

| $\bar{X}$      | $\leq 3.15$ | $3.15 - 3.45$ | $3.45 - 3.75$ | $3.75 - 4.05$ | $\geq 4.05$ |
|----------------|-------------|---------------|---------------|---------------|-------------|
| Observed Count | 20          | 20            | 28            | 21            | 11          |
| Expected Count | 25.84       | 20.47         | 21.52         | 16.75         | 15.42       |

(f) Since 0.229 is greater than 0.05 (or any commonly used significance level),  $H_0$  is not rejected.

The results of the investigation do not provide convincing evidence that the sampling distribution of  $\bar{X}$  is not exactly normal.

(g) The sampling distribution of  $\bar{X}$  is not exactly normal, so  $H_0$  is false. The correct result of the hypothesis test would therefore be to reject  $H_0$ . We have failed to reject  $H_0$  when  $H_0$  is false, and so this is a Type II error.

(The numerical results of this approach can be found by using a calculator, as follows. On the TI-83 and TI-84 calculators STAT TESTS, T-Test. For p-value, select the option  $\leq$  and then press ENTER, and if a statistic is relevant to the test being displayed.)

4. (B) The patients with no active ingredient in the Group B tablets are a placebo. If no placebo were given, and the Group A patients had better success than the Group B patients, then it would be impossible to know whether this effect was due to the drug or to the psychological effect experienced by the Group A patients of being given a treatment. However, if the Group B patients are given a placebo, then both groups experience this psychological effect, and any difference in the results between the two groups can more easily be attributed to the effects of the drug.

Options (A) is incorrect because it is based on the assumption that all patients respond to the drug with the same patients who do not take the drug. Option (C) is incorrect because only the patients in Group A receive the drug. Option (D) is incorrect because, though it is true that only the patients in Group B experience any side effects of the drug, this remains true whether or not Group B receives a placebo. Option (E) is incorrect because the purpose of giving a placebo is to blot out the psychological effect of taking a drug.

5.  $20.0 + 12.0 + 8.0 + 6.0 = 46.0$  (This is the total number of people in each treatment group). If the experimental were to be run on 6000 subjects, it is possible that the two groups would be statistically different in some way that was evident in the effectiveness of the drug. When the