A Sourcing Problem

RWW

10/14/2021

# Midterm Case: Help UNG Out

Answer the following questions **by yourself** in as much detail as possible. I take academic integrity very seriously and I know that you all do as well as it is what retains, or in its absence destroys, the value of our joint mission to excel. Partial credit will be awarded in the event that the reasoning is sufficiently clear to uncover the parts of a question correctly answered; be systematic in solving the problems. This will always take you far. Supply answers in relevant metrics with complete answers; computer output alone is insufficient. Faulty work could cost the company millions. Moreover, we really do wish to explain what we are discovering along the way to build toward a justification for our recommendation in the final phase.

## The Data

You have been given three data sets for use in tackling the problem. We have two data files that are identical: an Excel workbook and an .RData file.

* *Financials*: In one spreadsheet, we have revenues and expenses data for the last 36 weeks for two product lines in frozen foods (measured in 1000s of dollars per week). Pay no mind to the times that may appear alongside the dates this is an Excel vestige.
  + Tidy: *Revenues* and *Expenses* are recorded by *Product* by *Week* and *WeekNo*.
  + *Kiev.Expense* records Expenses for the Chicken Kiev product in 1000s of US dollars.
  + *Kiev.Revenue* records Revenues for the Chicken Kiev product in 1000s of US dollars.
  + *Marsala.Expense* records Expenses for the Chicken Marsala product in 1000s of US dollars.
  + *Marsala.Revenue* records Revenues for the Chicken Marsala product in 1000s of US dollars.
* *Satisfaction*: In one spreadsheet, the levels of product satisfaction are recorded for random samples of customers for each product.  
  + Tidy: *Response* records Like or NotLike for the given *Product* in that row.  
  + *Kiev.Like* records Like or NotLike for the sample asked about Chicken Kiev.  
  + *Marsala.Like* records Like or NotLike for the sample asked about Chicken Marsala.
* *InputCost*: In one spreadsheet, the costs of inputs for the two products are described in Raw and Tidy form.  
  + Tidy: *Cost* is the realized cost of inputs (per unit) in US dollars for *Product*  
  + *Kiev.Cost* is the realized cost of inputs (per unit) in US dollars for Chicken Kiev.  
  + *Marsala.Cost* is the realized cost of inputs (per unit) in US dollars for Chicken Marsala.

## The Problem:

A major grocer’s food division has a problem and your job is to recommend a course of action. UnNamed Grocer (UNG) is currently outsourcing the production of Chicken Kiev and Chicken Marsala to the same vendor. The vendor has suffered permanent setbacks due to COVID-19 related supply chain decimation and can no longer fully supply the grocer’s required quantities for both products. The vendor can continue to supply either one of the two products but not both and a new vendor will be needed if both products are continued. We will wish to examine the customer satisfaction, the financial elements, and cost structures are important elements in the decision because margins in the business are quite tight.

Key company procedures and definitions precede questions in some sections. This happens somewhat frequently so there are well established rules to use to evaluate the data in many cases. The company demands use of a 95% confidence interval/level for all decision making unless otherwise noted. Never forget that R’s help ? is your friend.

## [1] "Financials.Raw" "Financials.Tidy" "InputCost.Raw"   
## [4] "InputCost.Tidy" "Satisfaction.Raw" "Satisfaction.Tidy"

If we see the names of the spreadsheets above, they are active in this point in the Markdown. As described below, there are three spreadsheets that have been imported into six data.frames in R.

### The Basics

1. What is the probability
   * between -1 and 1 for a standard normal distribution?
   * between -1.5 and 0.8 for a standard normal distribution?
2. What values of a normal distribution with mean 25 and standard deviation 5 describe the middle 80% of the data? Provide both boundaries.
3. Suppose label misprints are described by a Poisson process and arrive at a rate of 6 per day, what is the probability of 10 or more misprints in a day?

*The log-normal probability distribution is a continuous probability distribution of a random variable whose logarithm has a normal distribution. If is log-normal, then is normal (or . It is also known as Galton’s distribution after Sir Francis Galton, a legend in the history of statistics, it is used to characterise how long people tend to linger on internet articles and the length of reddit posts along with the duration of games of chess and the size of living tissues. Other applications include hospitalizations during pandemics. In R, it is (d,p,q,r)lnorm(?, meanlog=0, sdlog=1) by default.*

1. What is the probability that a log-normal random variable [in R, lnorm is the noun] whose logarithm has mean equal to 0 and standard deviation equal to 1, takes values less than 1?
2. What is the value of a log-normal random variable [in R, lnorm is the noun] whose logarithm has mean equal to 0 and standard deviation equal to 1, such that 0.95 of the probability is below said value?
3. Plot a random sample of 1000 random draws from the aforementioned lognormal distribution (meanlog 0 and std. dev. 2).

### Input Costs

1. Provide summary statistics for each product and a combined graphic depicting the *Input.Cost* of Chicken Kiev and Chicken Marsala.
2. Suppose that each product has input costs that are normally distributed with the mean and standard deviation given in the summaries above (you may round to two digits to the right of the decimal place). What is the probability of costs above $2.25
   * for Chicken Kiev?
   * for Chicken Marsala?
3. What is the average input cost for Chicken Kiev with 95% confidence [provide both bounds]?
4. What is the probability that the average cost of Chicken Kiev is greater than $2.65?
5. What is the average input cost for Chicken Marsala with 95% confidence [provide both bounds]?
6. What is the probability that the average cost of Chicken Marsala is less than $2.45?
7. Does either product have lower input costs with 95% confidence?
   * If yes, which one?

### Customer opinions matter.

**Rule:**  
\* **The true probability of Like must be greater than or equal to 0.5 to continue a product.**

#### Questions

1. Provide a table showing opinions by product.
   * What is the probability of Like for Chicken Kiev?
   * What is the probability of Like for Chicken Marsala?
   * Which product is more popular, taking no account of uncertainty?
2. Management is concerned that the survey data may be influenced by missing responses. Each product was given to 130 tasters and free samples usually lead to probabilities of a response of 0.95. What is the probability of the number of observed responses [or fewer] for each product? **Less than 5% probability should lead you to discount the survey evidence though you must still answer the remaining questions**.
3. Should *Chicken.Marsala* be **continued**?
   * What is a (2-sided) 95% confidence interval for the probability of *Like*?
4. Should *Chicken.Kiev* be **continued**?
   * What is a (2-sided) 95% confidence interval for the probability of *Like*?
5. Is either product preferred with 95% confidence?
   * If so, which one?

### Product line financials

Product line financials are equally or more important. **Average net income is the key decision metric.** It should obviously be maximized. You can use R or Microsoft Excel for the calculation of net income.

**Rules:**  
\* **Products with average revenues less than 100,000 must be discontinued.**  
\* **Outstanding food products average over $10,000 in net income per week and must be continued.**  
\* **Products should be continued if their average net incomes are positive.**  
\* **Products should be prioritized if their average net incomes are different.**

## Basic Financial Variables

1. Provide a numerical summary of revenues and expenses for each product. Briefly discuss and compare the elements of your summary of choice.
2. Provide a graphic describing revenues by product.
3. Provide a graphic describing expenses by product.
4. Provide a 95% confidence interval for average revenues for each product.
5. Assume that *expenses* are distributed normal with mean and standard deviation exactly equal to those in the samples of data that you possess. *Provide a 90% [10% below, 90% above] lower bound for expenses for each product.*

### Net Income

1. Summarize net income for each product.
   * Provide a plot of Net Income over time.
2. Which, if any, products **must be discontinued**?
   * Provide 95% confidence intervals for average revenues for each product.
3. Which, if any, products **are outstanding**?
   1. What is the 95% confidence interval for net income for **Chicken.Kiev**?
   2. Should **Chicken.Kiev** be continued (by net income)?
   3. What is the 95% confidence interval for net income for **Chicken.Marsala**?
   4. Should **Chicken.Marsala** be continued (by net income)?
4. Is there a difference in net income between the two products?
   * Which one is more profitable?
   * Provide the [two-sided] 95% confidence interval for the average difference in net income.
5. What would you project average *Net Income* for **Chicken.Kiev** to be over an entire year?

## Some Other Concerns

Some recipes are more prone to harboring disease than others. The company is willing to accept the occasional risk but wishes to discontinue even moderately dangerous products.

Historically, Chicken Marsala has had recalls arrive at a rate of 4 per year. What is the probability of seven or fewer recalls in a given year for Chicken Marsala?

## The Choice [15%]

1. Discontinue both products
2. Continue only Chicken Marsala and cancel Chicken Kiev,
3. Continue only Chicken Kiev and cancel Chicken Marsala
4. Continue Chicken Kiev with the current vendor and seek a new vendor for Chicken Marsala.
5. Continue Chicken Marsala with the current vendor and seek a new vendor for Chicken Kiev.

Which do you recommend and why?

Suumarise all of the relevant findings as justification for up to 5 bonus points, summarize everything that is important in a single graphic.