**Outrigger Hotel Late Night Happy Hour**

A recent UH BUS310 student had a summer internship at a hotel in Waikiki. The first day on the job her boss handed her a spreadsheet with revenue for sales at the hotel bar for the past 24 weeks. He asked her look at the "variance" in the data to determine if adding a happy hour on 5/1/2013 (Week18) improved bar sales (in terms of revenue). The decision to be made is whether to keep the happy hour or stop offering it.

The student immediately thought of using a control chart. While there are other ways to answer this question, creating a control will reveal useful information. When interpreting this control, chart take care to look for the "story" rather than rule violations. This is an example where "out of control" implies an improvement to the process that you wish to maintain.

In this case you should make use of: hypothesis testing, qqplot (npp), unstacking, dummy variables

1. Specify the problem (goals, decisions, factors)

Goal: maximize profit from bar

Decision: Continue late night happy hour promotion?

Factors: expected total bar profit per day, impact of happy hour on profit

“Does late night HR increase profit?” 🡪 If yes then continue

1. Wrangle the data

* Appropriate format

Date – timestamp better?, Date: -> Date

Revenue time ranges – labels not convenient

fix(Late\_night\_happy\_hour)

Late\_night\_happy\_hour$Date <- as.Date(Late\_night\_happy\_hour$Date)

* Quality, integrity, representativeness

Time series by day – ok so long as enough before and after happy hour started

Date – label

Revenue by Time Ranges – Ratio, unequal time ranges? 🡪 use total (don’t really care about impact on revenue in particular time only overall impact)

\*\*\*Possible data insufficiency for answering the question: Profit requires cost, no cost data. Lower prices may increase demand thus revenue but with lower profit margin. Possible to have higher revenue but lower profit. Can we assume profit loss is approximately the difference in revenue HR/no-HR? Assume variable cost of drinks is not the primary cost (labor and space is dominant) and thus negligible effect on profit? Proceed with caution in interpreting results!

* New variables

TotalRev

Late\_night\_happy\_hour$TotalRev <- Late\_night\_happy\_hour$Rev9to10 + Late\_night\_happy\_hour$Rev10to11 + Late\_night\_happy\_hour$RevLateNight

Or

Late\_night\_happy\_hour$T <- apply(Late\_night\_happy\_hour[,c(2,3,4)],1,sum)

HappyHour – dummy variable

Late\_night\_happy\_hour$HR <- (Late\_night\_happy\_hour$Date > "2013-05-01")\*1

1. Models for the decision

* Characteristics, relationships, impact

1. “Does late night HR increase profit?” 🡪 is HR related to TotalRev?
2. Does HR affect revenue in non-HR times? Possible issue is people wait for HR and thus lower higher profit making non-HR revenue to the point where overall profit is less.

Look boxplots of TotalRev to compare groups and see if there are differences (essentially comparing distributions). Why is a time series plot not useful here?

boxplot(Late\_night\_happy\_hour$TotalRev ~ Late\_night\_happy\_hour$HR)



Clearly no HR TotalRev is less, variability about the same, some outliers,

Mu\_0 = pop. Mean TotalRev with no HR

Mu\_1 = pop. Mean TotalRev with HR

H0: Mu\_0 = Mu\_1

H1: Mu\_0 <> Mu\_1

totalRev\_anova <- aov(Late\_night\_happy\_hour$TotalRev ~ Late\_night\_happy\_hour$HR)

summary(totalRev\_anova)

Df Sum Sq Mean Sq F value Pr(>F)

Late\_night\_happy\_hour$HR 1 256974 256974 11.95 0.000692 \*\*\*

Residuals 168 3612275 21502

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Can be nearly 100% confident that TotalRev is higher with HR

anova$coefficients

(Intercept) Late\_night\_happy\_hour$HR

359.16759 85.83772

Impact expect (on average) $85.83 more TotalRev when HR (note this is same as SLR).

Could also look at t-test

H0: Mu\_0 - Mu\_1 = 0

H1: Mu\_0 - Mu\_1 <> 0

t.test(TotalRev ~ HR, data=Late\_night\_happy\_hour)

Welch Two Sample t-test

data: TotalRev by HR

t = -3.3223, df = 81.96, p-value = 0.001335

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-137.23570 -34.43974

sample estimates:

mean in group 0 mean in group 1

359.1676 445.0053

With 95% conf. we expect (average) TotalRev is between $34 and $137 more revenue with HR

Late\_night\_happy\_hour$TotalRevEarly <- Late\_night\_happy\_hour$Rev9to10 + Late\_night\_happy\_hour$Rev10to11

boxplot(Late\_night\_happy\_hour$TotalRevEarly ~ Late\_night\_happy\_hour$HR)



summary(aov(Late\_night\_happy\_hour$TotalRevEarly ~ Late\_night\_happy\_hour$HR))

Df Sum Sq Mean Sq F value Pr(>F)

Late\_night\_happy\_hour$HR 1 82327 82327 6.122 0.0143 \*

Residuals 168 2259388 13449

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

aov\_early$coefficients

(Intercept) Late\_night\_happy\_hour$HR

277.28825 48.58522

Can be over 99% conf. that Early Rev with HR > Early Rev no HR. Expect about $49 more.

boxplot(Late\_night\_happy\_hour$RevLateNight ~ Late\_night\_happy\_hour$HR)

summary(aov(Late\_night\_happy\_hour$RevLateNight ~ Late\_night\_happy\_hour$HR))

Df Sum Sq Mean Sq F value Pr(>F)

Late\_night\_happy\_hour$HR 1 48400 48400 6.744 0.0102 \*

Residuals 168 1205713 7177

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

aov\_late$coefficients

(Intercept) Late\_night\_happy\_hour$HR

81.87934 37.25250

Can be over 99% conf. that Early Rev with HR > Early Rev no HR. Expect about $37 more.



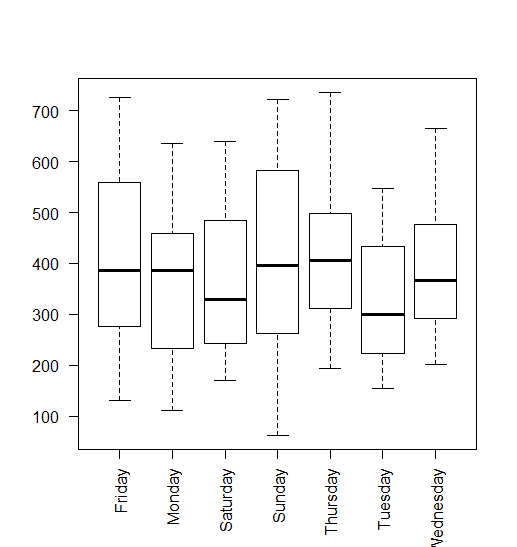
* List and check assumptions

Check Normality and variance of smaller groups

Does day of the week matter?

L$Day <- weekdays(as.Date(L$Date))

boxplot(L$Total ~ L$Day, las=2)



days\_anova <- aov(L$Total ~ L$Day)

summary(days\_anova)

Df Sum Sq Mean Sq F value Pr(>F)

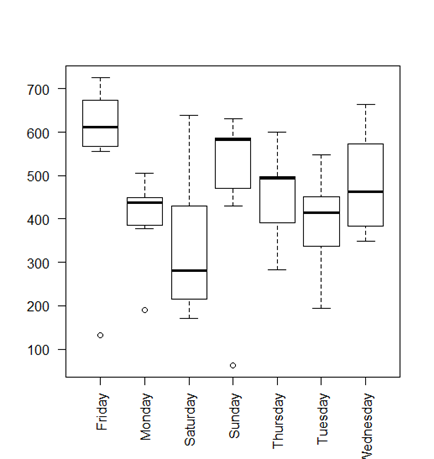
L$Day 6 168530 28088 1.237 0.29

Residuals 163 3700718 22704

Day of week doesn’t seem to matter notably overall, but maybe for when there is a happy hour it does?

WHR <- subset(L, HR==TRUE)

NHR <- subset(L, HR==FALSE)



boxplot(WHR$Total ~ WHR$Day, las=2)

Friday seems to be substantially higher than other days and Saturday mildly lower.

summary(aov(WHR$Total ~ WHR$Day))

Df Sum Sq Mean Sq F value Pr(>F)

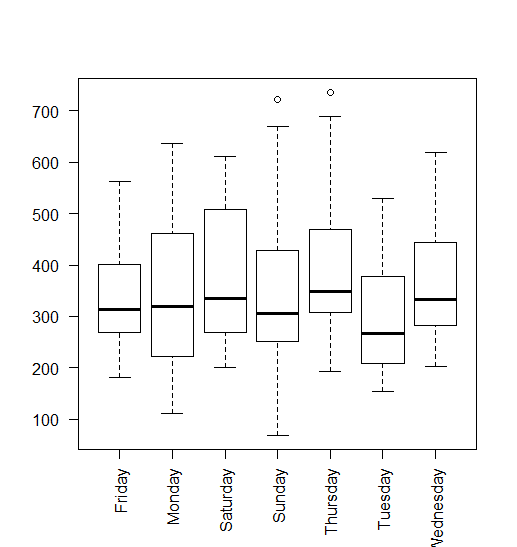
WHR$Day 6 234461 39077 1.744 0.134

Residuals 42 940844 22401

Hmm… can only be about 86.6% confidence that some means are different. Maybe there’s not as much of an effect as it seems.

What about no happy hour?

> boxplot(NHR$Total ~ NHR$Day, las=2)



Nothing to see here!

qqplot

Time of year matter? Organize by month and check for differences.

Revenue proxy for profit? 🡪 check drink material costs

1. Interpret results

So what is the specific advice for the decision here and why (backed up by the analysis and data)?