

Statistical Analysis

Barry Davis

September 6, 2017

```
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)

## Warning: package 'dplyr' was built under R version 3.4.1
library(lubridate)

## Warning: package 'lubridate' was built under R version 3.4.1
library(data.table)
library(plotly)

## Warning: package 'plotly' was built under R version 3.4.1
answered <- fread("answered_primary_refined.csv", na.strings = c("", "NA"))
aperf <- fread("agent_perf_refined.csv", na.strings = c("", "NA", "\\N"))
```

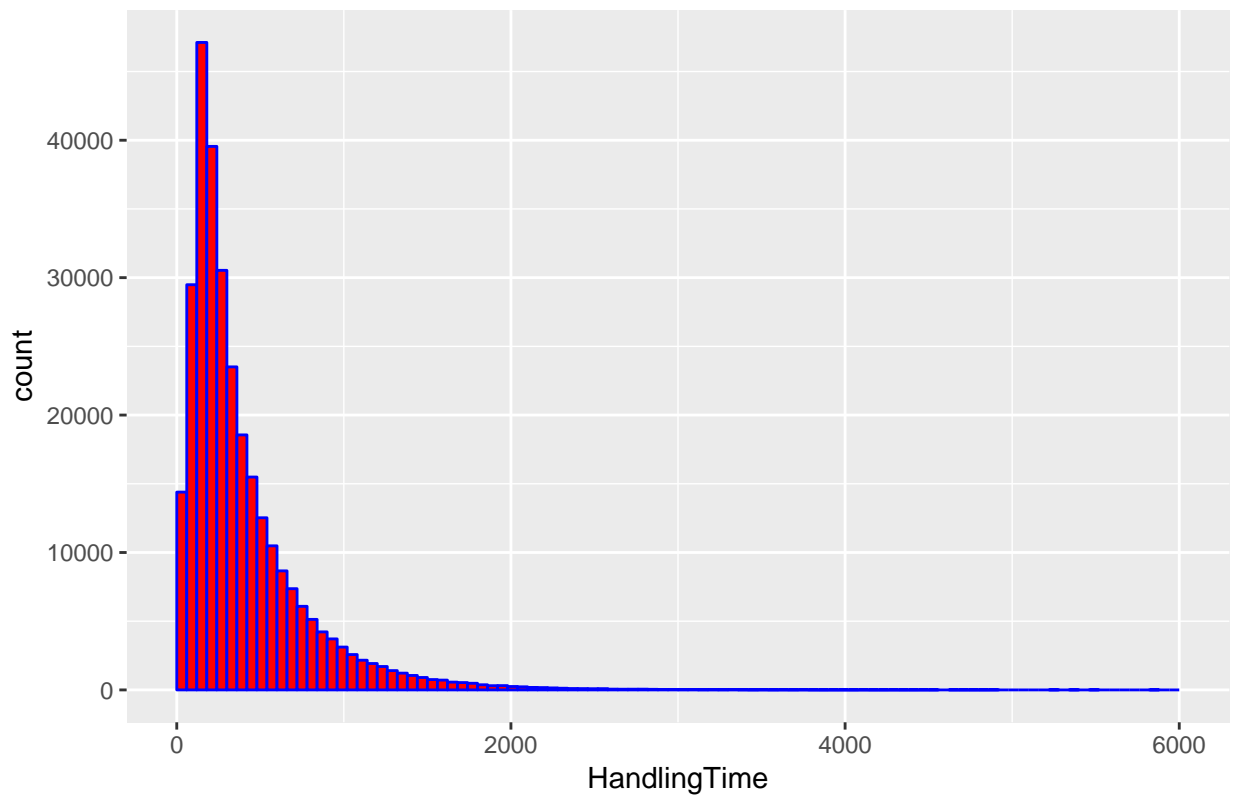
Introduction

This document will review some statistical analysis of the call center data for agents. Since my main objective has covered some of these ideas regarding call volume, I will focus on agent data for this analysis. Agent behavior directly affects call queue wait times.

Handling Time

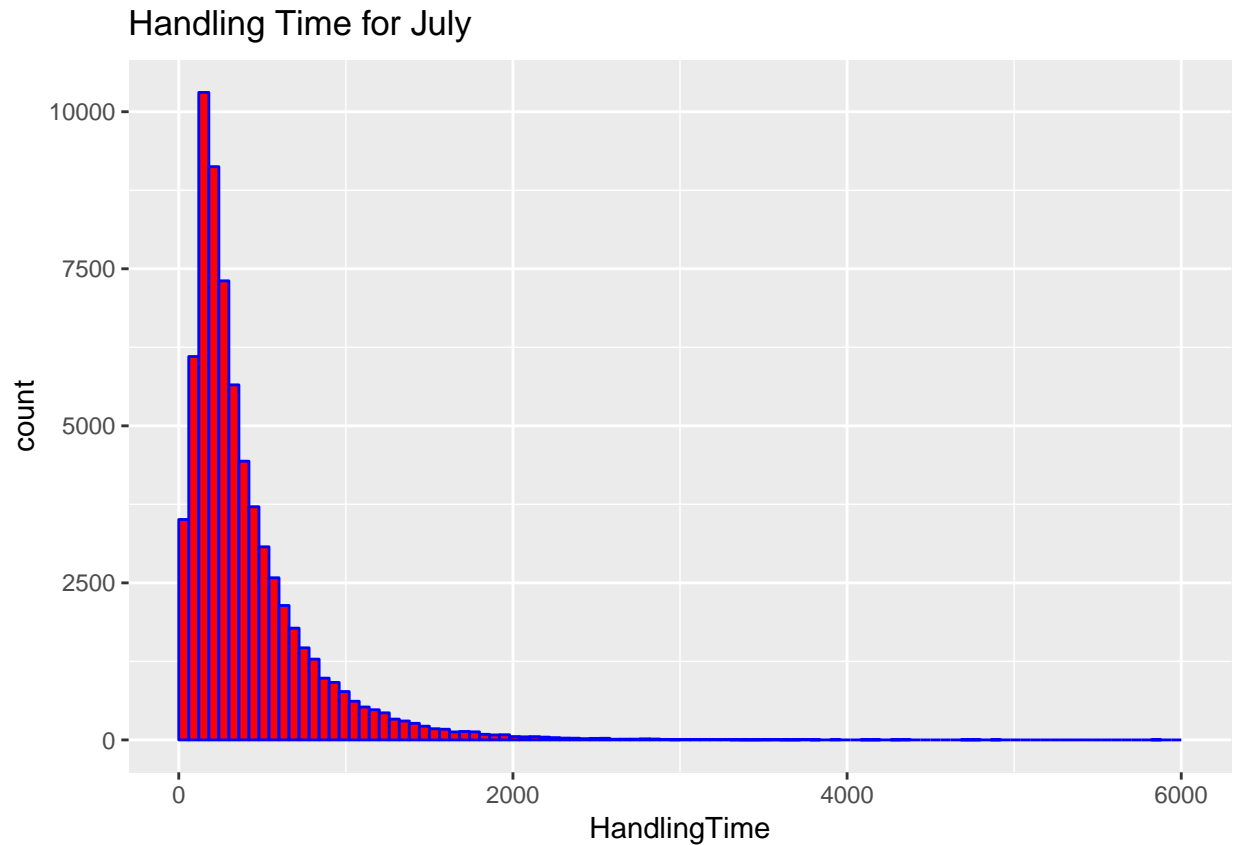
```
ggplot(answered, aes(HandlingTime)) +
  geom_histogram(breaks = seq(0, 6000, by = 60), col="blue", fill="red") +
  labs(title="Handling Time for 5 months")
```

Handling Time for 5 months



Here we see the agent handling time of a call. Most of the calls are falling into the 0-7 minute range, with quite a long tail extending out to a max of 5868. Perhaps it would be better to view these in a given month rather than all together.

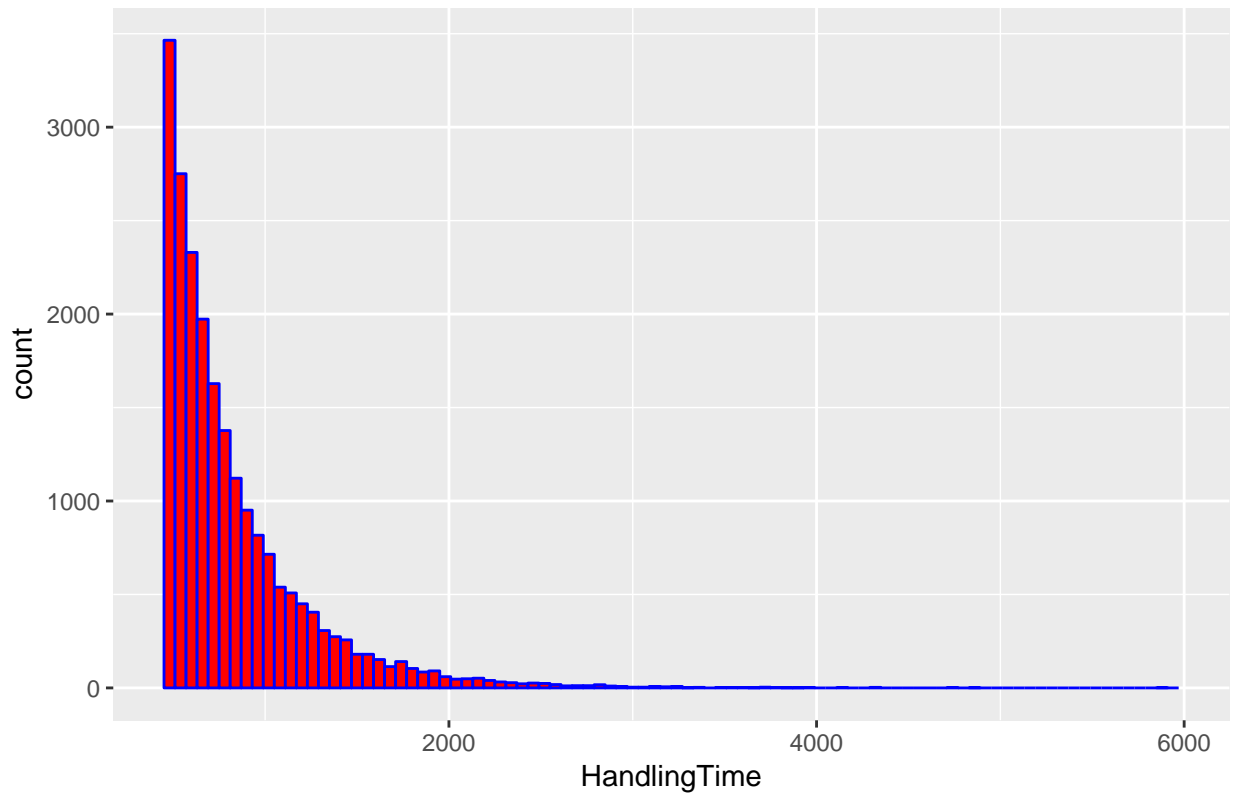
```
answered %>% filter(month(OriginatedStamp) == 7) %>% ggplot(aes(HandlingTime)) +  
  geom_histogram(breaks = seq(0, 6000, by = 60), col="blue", fill="red") +  
  labs(title="Handling Time for July")
```



While looking at the same data for July, we can see that the pattern approximately follows the same trend we see for our plot of all our data. There are a number of calls that go well beyond the targeted handle time of 330 seconds and allowing a range up to 450 seconds. So let's look at those in particular.

```
answered %>% filter(month(OriginatedStamp) == 7) %>% ggplot(aes(HandlingTime)) +  
  geom_histogram(breaks = seq(450, 6000, by = 60), col="blue", fill="red") +  
  labs(title = "Calls over 450 seconds in July")
```

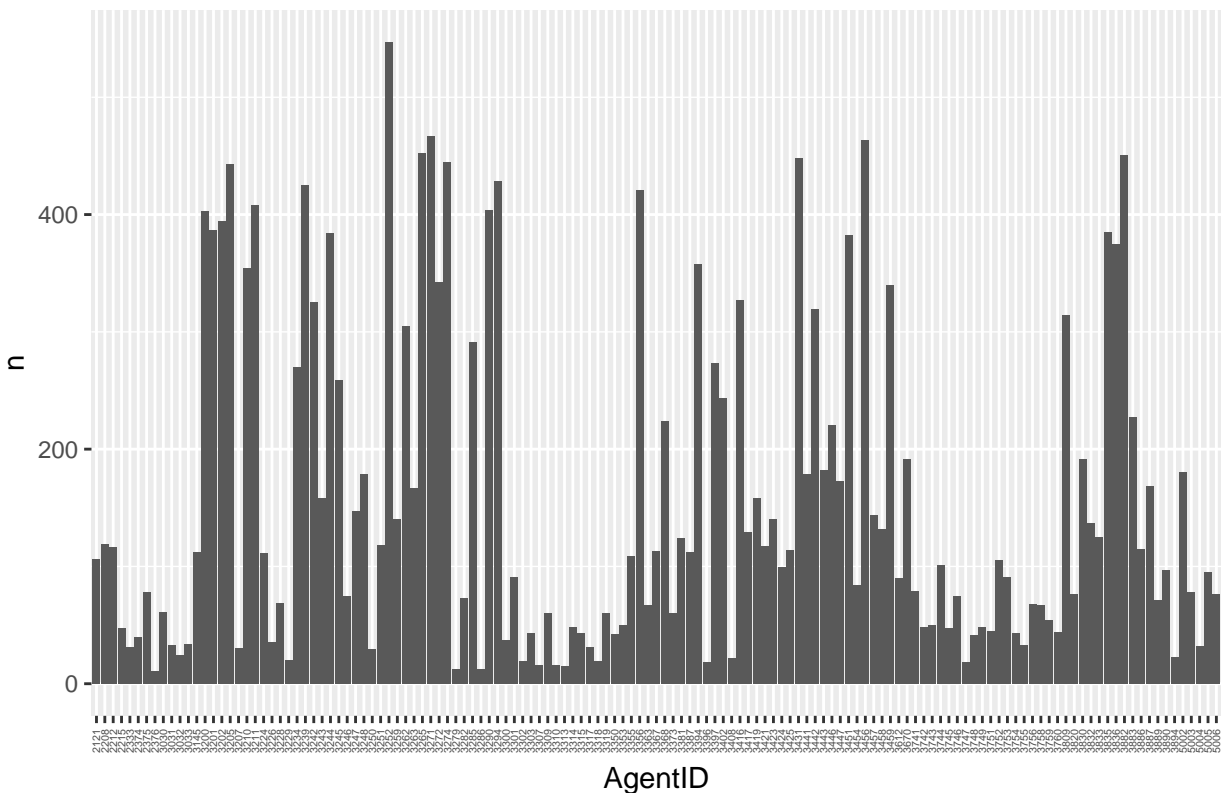
Calls over 450 seconds in July



All of these are extending the wait times of customers. These could be an indication that some agents aren't able to handle calls appropriately, but due to the nature of a call center, it very well could be the luck of the draw on which agents gets a call that will be long regardless of the agent who handled it. I will plot this data by agent to see if there's any repeat offenders.

```
answered$AgentID <- answered$AgentID %>% as.factor()
agents_over_450_sec_july <- answered %>% filter(HandlingTime > 450, month(OriginatedStamp) == 7) %>% count(AgentID)
agents_over_450_sec_july %>% ggplot(aes(x = AgentID, y = n)) +
  geom_col() + labs(title = "July Agents with 10+ Calls Longer Than 7.5 Minutes") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5, size = 4))
```

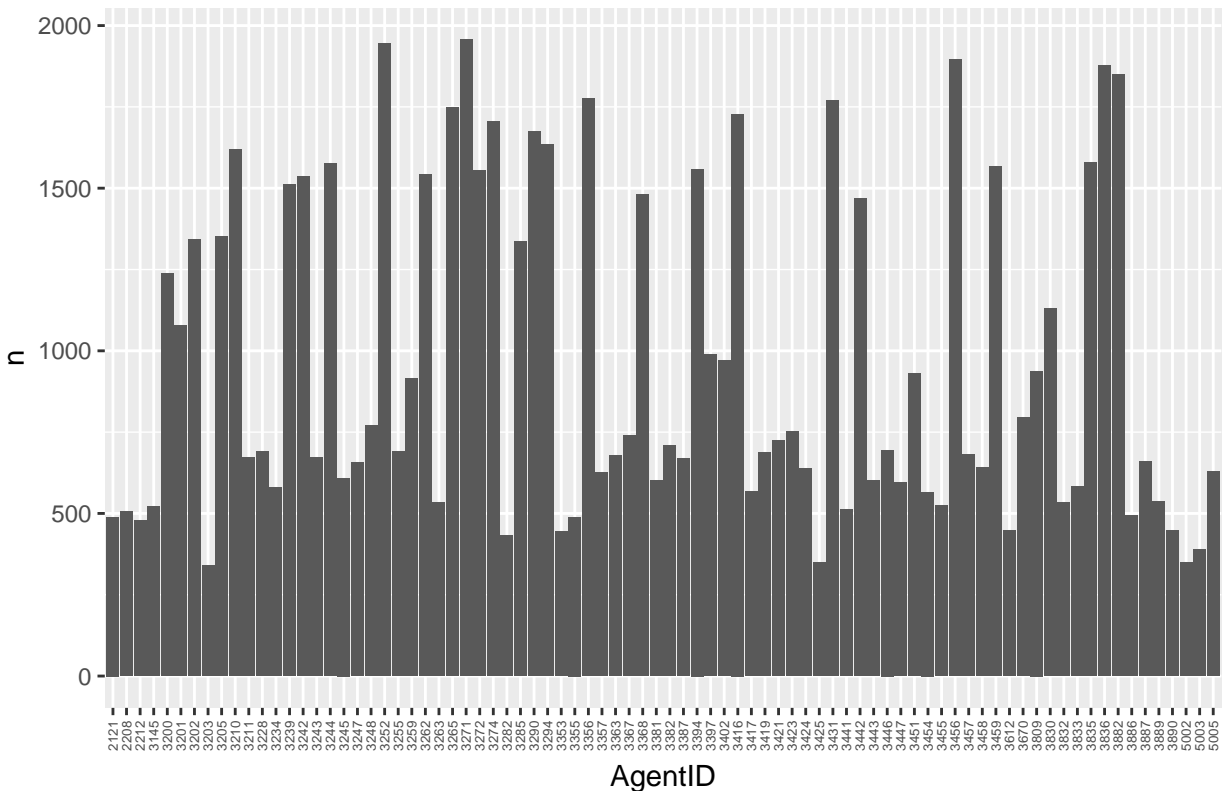
July Agents with 10+ Calls Longer Than 7.5 Minutes



This offers some nice insight. This has been filtered to limit to agents who have more than 10 calls which are above 450 seconds (7.5 minutes). These are the top trouble calls causing the most additional wait time for customers. Next I'd like to repeat this filter on our total data set. I'm also going to set a threshold of 480+ calls. The data starts on 3/30/17 and ends on 8/20/17, and we should give a little leeway. So, on the generous side, if an agent worked every day and received 2 calls per day over 10 minutes in length each, that would give each agent roughly 60 calls per month. So, 300 calls would cover 5 months of bad luck calls.

```
agents_over_450_sec <- answered %>% filter(HandlingTime > 450) %>% count(AgentID) %>% filter(n > 300)
agents_over_450_sec %>% ggplot(aes(x = AgentID, y = n)) +
  geom_col() +
  labs(title = "Agents with 300+ Calls of 7.5 Minutes or Longer in 5 Months") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5, size=5))
```

Agents with 300+ Calls of 7.5 Minutes or Longer in 5 Months



```
num_calls_over_450_sec <- agents_over_450_sec %>% tally(n)
seconds_over_450_sec <- answered %>% filter(HandlingTime > 450) %>% tally(HandlingTime)
```

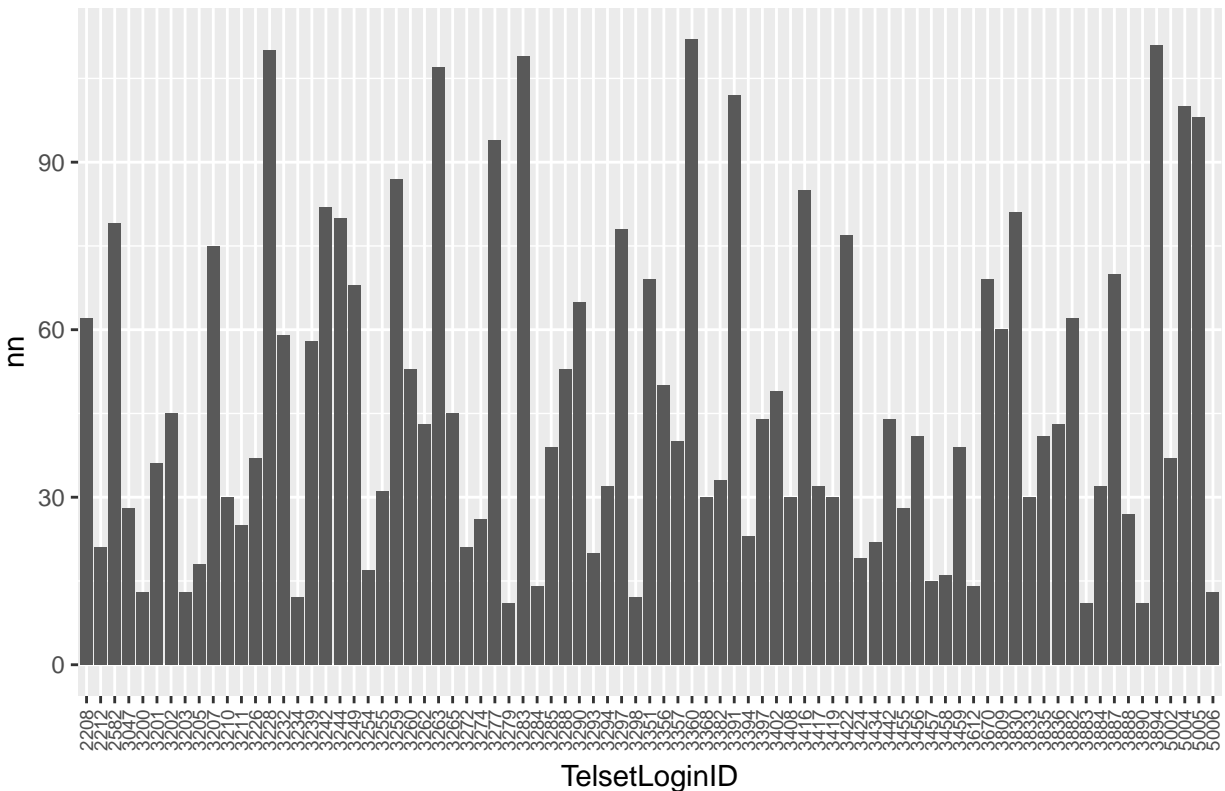
It looks like these agents may need additional training. In total, that's 78765 calls in 5 months, each over 7.5 minutes in length. That's a total talk time of 2.0219821×10^4 hours or an average of 15.4 minutes per call that's over 7.5 minutes. There may be an opportunity here for the call center to reduce the handling time with these agents, which would also increase the number of calls answered.

Not Ready Time

Another important statistic in call centers is Not Ready Time. This is essentially time that an agent is logged into the phone, but is in a "not ready" state. During this time, no new calls will be presented to the agent. Every minute spent in a not ready state is a direct impact to the wait time for customers in queue. Therefore, we should look at this data point as well to determine if there are any issues that should be addressed.

```
aperf$TelsetLoginID <- aperf$TelsetLoginID %>% as.factor()
aperf <- aperf %>% separate(statTimestamp, c("Date", "Time"), sep=" ", remove=FALSE)
agents_not_ready <- aperf %>% select(Date, TelsetLoginID, NotReadyTime) %>% group_by(Date, TelsetLoginID)
agents_not_ready_over_1440 <- agents_not_ready %>% ungroup() %>% select(TelsetLoginID, n) %>% count(TelsetLoginID)
agents_not_ready_over_1440 %>% ggplot(aes(x = TelsetLoginID, y=n)) +
  geom_col() +
  labs(title="Agents 10+ days of over 24 minutes Not Ready in 6 months") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5, size = 7))
```

Agents 10+ days of over 24 minutes Not Ready in 6 months



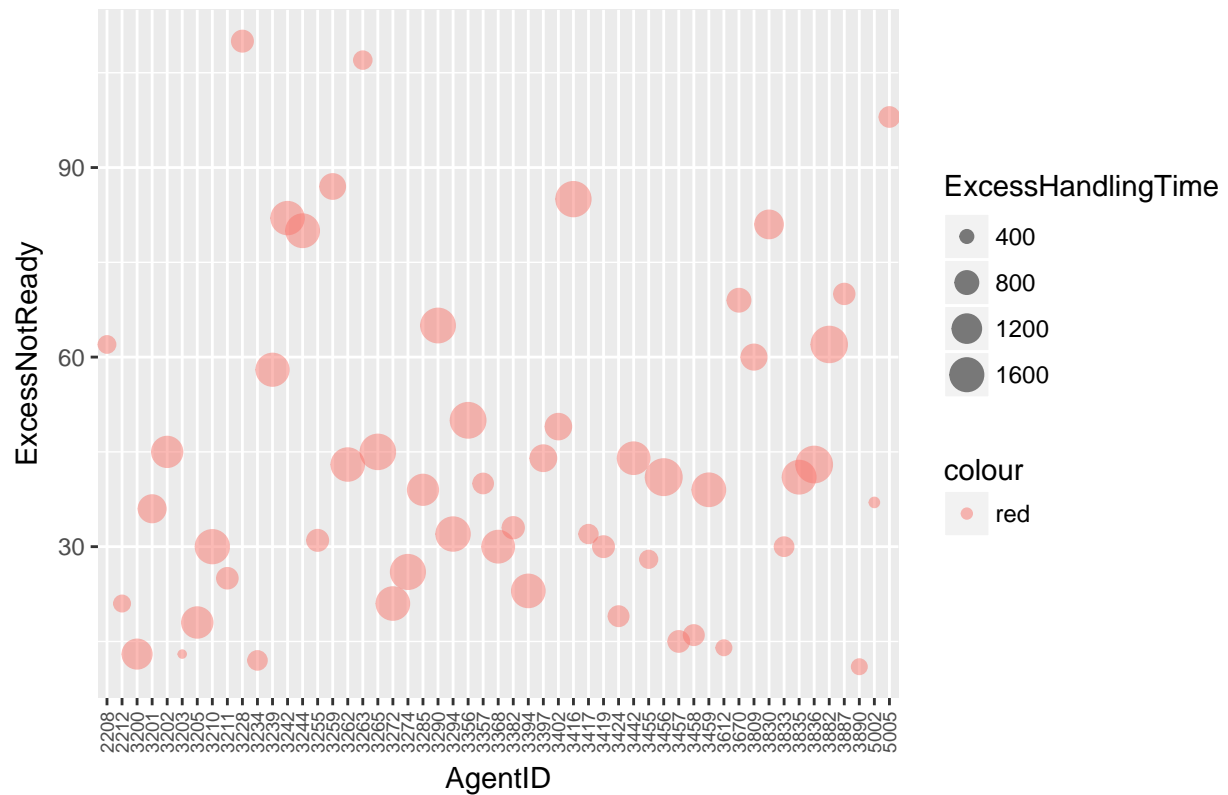
As we can see, there are a number of agents who are going beyond the allotted 24 minutes of Not Ready time. Reducing these instances will also improve queue times for customers. These agents likely could use refreshers on appropriate time management. The reason for being Not Ready could vary, so each case may need to be handled individually.

Combined Agents

I thought it would be interesting to see which agents are in both groups. This would mean that the agent had more than 300 calls that were over 450 seconds in duration, as well as 10 or more days with a total Not Ready time of more than 24 minutes.

```
trouble_agents <- merge(agents_not_ready_over_1440, agents_over_450_sec, by.x = c("TelsetLoginID"), by.y = c("AgentID"),
  setnames(trouble_agents, old = c('nn', 'n'), new = c('ExcessNotReady', 'ExcessHandlingTime'))
trouble_agents %>% ggplot(aes(x = TelsetLoginID, y = ExcessNotReady, col = "red")) +
  geom_point(aes(size=ExcessHandlingTime), alpha=.5) +
  labs(title="300+ calls over 7.5 minutes and 10+ days of Not Ready time over 24 minutes", x="AgentID")
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5, size=7))
```

300+ calls over 7.5 minutes and 10+ days of Not Ready time over 24 minutes



Conclusion

Before spending additional money on salaries and benefits for new hires, it may be worth ensuring the existing employee base is meeting expectations on Handling Time and Not Ready Time. Reducing both of these will have a positive impact on customer queue times.