Crowds and Corona:

The influence of covid-19 on a music festival

Group 3
Sofia Fraile Zendejas and Jeanine Buurma

Introduction

Crowd-based events are generating new forms of crowd-based performative interaction.(Sheridan et al., 2011) The way crowds behave in large-scale events is an interesting field to research that could potentially bring new opportunities for the planning of the events as well as crowd management and increased user experience from the festival go-ers.
This project is based on crowd behaviour analysis for music festivals and it is relevant for the modelling of large-scale events to contribute to crowd management, that leads to avoid accidents as well as a better user experience for the attendees, as the modelling of these circumstances can be used in both perspectives to create an overall better festival experience.
A clear distinction between crowd control and crowd management should be made, since crowd management is the method to control a crowd in a normal situation and crowd control is the method to control a crowd in an emergency situation (Martella et al., 2017). An example of why crowd control is so important can be found in the Love Paradise disaster that took place on the 24th of july 2010. 21 people died and many were injured because of poor crowd control at this festival, which led to the attendees being crushed and getting into a state of panic (Klüpfel, 2014). Klüpfel shows that there are important feedback loops that can create these bottlenecks of busyness and eventual pushing and chaos. Important factors that should be well managed during an event that are mentioned are time of opening of the event, unexpected counterflow, insufficient management of the inflow and non existing alternative stages.
Both crowd control and crowd management are very important but we will focus our model on crowd management.

2020 and 2021 have introduced a new form of crowd management because of the ongoing pandemic. There are a large number of guidelines that are written for places with a large number of people yet a lot of scenarios have not been tested in real life yet because of the severity of the pandemic as of january 2021. The arrival of vaccines and our growing understanding of covid-19 could have a positive impact on the festival season in 2021, provided that they do follow some of the guidelines for crowd control during a pandemic. This could create some problems that normally would not arise as quickly, such as bottlenecks and locations/stages reaching their maximum capacity more quickly. Additionally not every attendee would always comply with all the guidelines especially at a music festival where the consummation of alcohol or other substances could make people less inclined to follow the guidelines. This also raises the question on what to do with repeating offenders of these guidelines. The enforcers of these guidelines also have to keep in mind the privacy of the attendees, which is a factor with normal scenarios as well (Pouw et al., 2020).

Having the two scenarios, crowd regulation with and without covid-19, in mind we haven chosen to the research the following:

Research questions:

1. How much does the average amount of waiting increase with social distancing in comparison to without?
2. Will wearing a mask have an impact on the speed the disease spreads?

To resolve these research questions, the following hypotheses were formulated:

1. Does the average amount of waiting increase with social distancing in comparison to without?
   - H_0: The average waiting amount does not increase with social distancing vs not social distancing
   - H_A: The average waiting amount increases with social distancing vs not social distancing
2. Will wearing a mask and social distancing have an impact on the speed the disease spreads?
   - H_0: Wearing a mask and social distancing has no impact in the amount of people infected
   - H_A: Wearing a mask and social distancing has an impact on the amount of people infected

Model description

Most models of crowds at festivals are based on common knowledge, previous experience, and trial and error (Martella et al., 2017) so we based our model on the same principles. Complete models of festivals and other large crowds often include transportation to location and security but we will focus our model on steering the crowd and the movement of a crowd within a festival terrain. This is why we exclude certain factors such as personnel, weather, institutions, emergencies, personal circumstances of attendees, toilets, camping grounds, emergency exits etc. The factors we do include are listed and described below.
The second version of the model has included covid-19 guidelines, most importantly the social distancing rule and the attendees wearing masks. This factor has influence on the distribution of the attendees and the speed they can move around terrain since their movement will be influenced by keeping distance from other attendees.

A general description of our model is that we have a festival terrain with multiple stages of different sizes, a bar and a relaxation area that are connected via paths. We have two types of turtles: attendees and artists. Artists can only go onto the stages to perform and attendees are free to go to wherever they want. A more in depth explanation per element in our model can be found below.

Attendees

Attendees are the core element of our model. Every attendee has random factors assigned to them to determine their action within the model. The first factor is how into music this attendee is which is expressed through a number between 1 and 5, with 5 being very into music and 1 not being interested in music at all. When an artist arrives on stage the sum of the artist's popularity and the music interest of an attendee must be a certain number to have the attendee attend the show of the artist. This is determined with a slider although a good threshold value is 6. The threshold-attendees slider can be seen as how important the performances are at a music festival with a low threshold determining that people attend the festival with seeing the artists as goal whilst a higher threshold determines that people attend the festival to party regardless of the artists performing. When there are multiple artists within an attendees threshold performing they will be randomly assigned to one of the performances. Another factor assigned to each attendee is their energy level which has a value between 1 and 10. Every attendee starts with an energy level of 10 and everytime they go to a performance this level

decreases by one. This level can increase by going to the bar or relaxation area, increasing the level by 2 and 1 respectively. When an attendee's energy level reaches 0 they are randomly assigned either the bar or relaxation area as a goal regardless of whether there is an artist they want to see performing or not. They can leave this area when their energy level is fully replenished.
Depended on their location, energy level and music preference they can have one of the following states with the default status being "idle":

- "idle"
- "going to performance"
- "choosing place at performance"
- "listening performance"
- "bar"
- "drinking"
- "relax"
- "do relax"

For the Covid-19 implementations, the following status are set:

- "Susceptible"
- "Infected"
- "Wearing a mask?"

The susceptible status, symbolized in green, applies to all attendees in a ratio of an infected person. Infected people show in red and people wearing a mask show in blue. In order for the infection to happen, we let healthy people in the radius of infection to be susceptible, and we provide mask wearers a random-float that will then need to be less than our threshold set by the probability of spreading.

Artists

Artists are the second type of turtle and have one factor that defines them, their popularity. Their popularity, or how famous they are, is a random number between 1 and 5 with 1 being very unknown and 5 being very famous as for example the Beatles or Ariana Grande. The higher their popularity score, the more likely attendees are to attend their concert as explained earlier. Artists do not have the same movement freedom as attendees and are only allowed on a stage. Artist can be in one of the following states:

- "idle"
- "performance"

Locations

There are four types of locations in our model, paths, stages, relaxation area and the bar. Paths do not have any function other than being the patches the attendee turtles use to get from location to location. The path patches are green.
The second locations are the stages, which have the components of the actual stage and the area where attendees can watch the stage. Attendees are not allowed on the stage and artists are not allowed off the stage. As described earlier, everytime a attendee enters a stage location 1 energy will be drained from their energy level every tick. The stage locations have different sizes, which becomes especially important when covid guidelines are applied. The stage patches are orange.
The third location is the relaxation area. In order to replenish energy attendees need a place to calm down after all the excitement of the performances. Attendees gain 1 energy level

back for every tick they spend inside of the relax area. The relax area is also the only place where attendees begin their festival “day”. The relax area patches are the color pink.
Lastly we have the bar area, whose patches are the colour purple. If an attendee chooses to go to this area when their energy level is 0, they gain 2 energy levels per tick they stay there. We have chosen to do an increase of 2 instead of 1 because a drink, some water or a soda can really give a person a pick-me-up when they feel tired.

Virus guidelines

There are multiple independent factors that can influence the spread and existence of a virus in our model. The r-distance is the distance in patches that turtles will keep between each other when they need to socially distance, which can have a value from 0.0 to 3.0. Whether they will socially distance can be decided with a switch and their distance is decided by a slider. Because of this distance there is a higher chance that bottle necks will be created. Realistically, people do not want to wait in line for a long amount of time and will probably break covid guidelines when annoyed with the waiting time. How long an attendee will wait in line can be decided with the max-patience slider which edits every attendees max waiting time in ticks.
Attendees can also wear masks and they turn blue in the model when they do so. How many of the attendees wear masks can be decided with a slider in percentages and the effectiveness of the masks can be edited with a slider as well.
Lastly the virus itself can also be edited. The probability of the spreading within a certain radius can be changed with a slider called “probability-spreading”. The radius in patches can also be changed with the slider radius-infection. In order to start the spreading of the virus one can flip the designated switch which will give one random attendee, who will be coloured red, the virus upon activation.

Experiments

To answer our research questions the following experiments were designed:

Experiment 1

For the research question: *Does the average amount of waiting time change when social distancing depending on the patience of the attendee ?* We established the following null and alternative hypotheses:

- H_0: The average waiting amount is not affected depending on patience level of the attendee
- H_A: The average waiting amount is affected depending on the patience level of the attendee.

With these hypotheses the independent variable patience-levels and the dependent variable waiting-time can be distinguished, taking into account the confounding variable of social distancing as true. For this, the experiment was set with the following variables:

["distancing?" true]
["threshold-attendance" 6]
["numberOfPeople" 121]
["max-patience" 50 100]
["r-distance" 1.5]
For 3 repetitions and time limit set to 1000 ticks.

This experiment will measure the waiting time that is set as a counter for the amount of times a person has to wait to move to a different location if abiding by the social distancing rules. The r-distance was set to 1.5 patches to simulate the 1.5 meter rule we have now in our society. Taking into account

different levels of patience that the person could have, the max-patience was set to 50 and 100, which are meant to symbolize the amount of time in ticks a person will wait before breaking the social distancing rules and moving towards the target patch.
This experiment is designed with the above mentioned assumptions as a way to isolate and get data only on the variables we are interested in.

Experiment 2

For the research question *Will wearing a mask have an impact on the speed the disease spreads?* We established the following null and alternative hypotheses:

- H_0: Wearing a mask has no impact on the amount of people infected
- H_A: Wearing a mask has an impact on the amount of people infected

With these hypotheses the independent variables of mask wearing can be distinguished. Mask wearing percentages will vary (50, 100) and social distancing will stay true at all times, this decision was made based on the assumption that people that are willing to wear a mask are also willing to accept social distancing. The dependent variable will be the amount of people infected taking into account the confounding variable of the probability of infection that was set to 0.5 arbitrarily for the purpose of this experiment as a way to isolate our variables. For this, the experiment was set with the following netlogo variables:

["distancing?" true] %social distancing on
["percent-wearing-masks" 50 100] %percentage of people wearing a mask
["threshold-attendance" 6]
["numberOfPeople" 121]
["spread-decease?" true] %1 person with covid
["probability-spreading" 0.5]
["max-patience" 100]
["r-distance" 1.5]
For 3 repetitions and a time limit set to 1000 ticks.

For this experiment the main goal was to measure the amount of people infected with the virus. The research question calls to be examined by setting different percentages of people wearing a mask, in this case 50% and 100%, in a set amount of time (1000 ticks).

RESULTS

Experiment 1

In order to attempt to answer the research question: *Does the average amount of waiting time change when social distancing depending on the patience of the attendee?* And the hypotheses:
H_0: The average waiting amount is not affected depending on patience level of the attendee
H_A: The average waiting amount is affected depending on the patience level of the attendee.
At first glance we can say that a t-test can be used for the data, however, to make sure the test is relevant and robust, the main requirement is that our data is normally distributed.
The first step taken was analysing the lack of extreme outliers that could potentially modify the distribution of our data, the function identify_outliers() from the rstatix module in R-studio, in the same fashion, a Shapiro test was used to verify the distribution of our data. Shapiro test for distribution takes the null hypothesis of data being normally distributed and an alternate hypothesis that the data is not normally distributed.
Which provides us the following plot (figure 1), where we can see that the data is heavy-tailed and therefore not normally distributed. A result for the Shapiro test the following values were obtained, W = 0.79466, p-value < 2.2e-16. This P-value is less than 0.05 (alpha value) which also provides evidence that the data is not normally distributed.

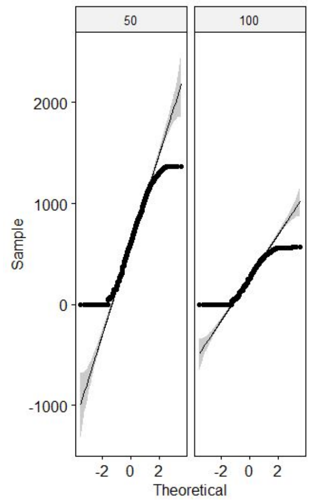

Figure 1

Since the data is not normally distributed, a Kruskal test should be used as it is a non-parametric alternative to the t-test, when the date does not match the assumptions for an ANOVA. To conduct this the following summary was created:

Patience level	variable	n	median	iqr	mean	SD	SE	CI
50	waitingTime	3003	618	598	619.804	372.366	6.795	13.323
100	waitingTime	3003	258	285	264.267	171.799	3.135	6.147

Where it can be observed that the median waiting time for attendees with a patience level of 50 ticks is 618(IQR = 598), and the median waiting time for attendees with a patience level of 100 ticks is then 258(IQR =285).

The test (code found in appendix)provides us with the following values:

Variable	group1	group2	n1	n2	statistic	p
waitingTime	50	100	3003	3003	7061140	2.2e-16

Variable	group1	group2	n1	n2	effectSize	magnitude
waitingTime	50	100	3003	3003	0.4902433	moderate

A moderately significant effect size of 0.490 and $p < \alpha$ (0.05)provides us with significant evidence to reject the null hypothesis that *the average waiting amount is not affected depending on patience level of the attendee.*

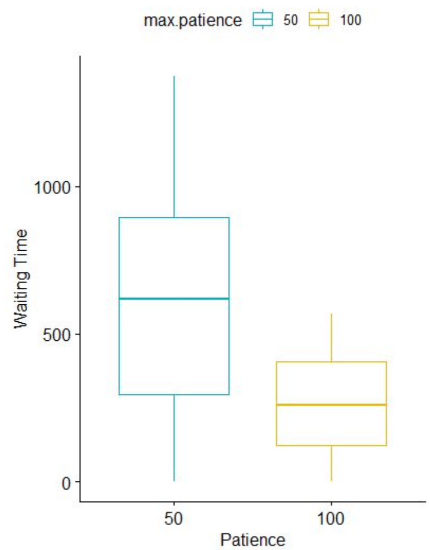

Figure 2

Experiment 2

For the second research question: *Will wearing a mask have an impact on the speed the disease spreads?* With the null and alternative hypotheses:
H_0: Wearing a mask has no impact on the amount of people infected
H_A: Wearing a mask has an impact on the amount of people infected.
The data needs to be analysed first to determine which test is suitable. We have used the test for normality of Shapiro to determine if our data is normally distributed or not, therefore assuming that the null hypothesis of the test says the data is normally distributed and our alternate says it is not. For that matter, we then encounter that the values for the samples (50, 100 percentage of masks being worn) are W = 0.79025, p-value < 2.2e-16, which states our p-value below our alpha value of 0.05. Therefore we can reject the null hypothesis and establish that the data is not normally distributed, which can also be verified with the following visual representation, that shows the heaviness of the tails.

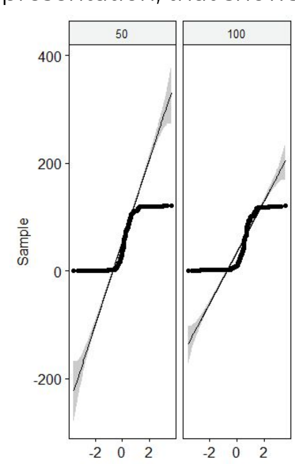

Figure 3

With this research question there was also an analysis of extreme outliers that could potentially affect the data, the function identify_outliers() from the rstatix module in R-studio, was also used for this experiment.
For data that is not normally distributed, a t-test cannot be performed, therefore the non-parametric alternative test that can solve the question is the Kruskal Wallis test, as it is an alternative to the one-way ANOVA test, that cannot be used since the assumption of having the independence of observations cannot be met.

The following summary of information was obtained:

Percent wearing a mask	n	median	IQR	mean	SD	SE	CI
50	2818	49	109	57.489	48.345	0.911	1.786
100	2944	3	66	33.663	44.769	0.825	1.618

This can be interpreted then as a median of 49 (IQR=109) when 50% percent of attendees wear a mask, and a median of 3(IQR=66) when 100% of attendees wear a mask.
The box plot can be observed here:

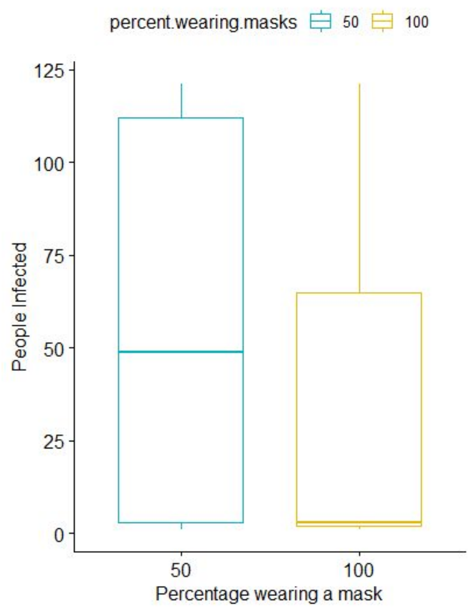

Figure 4

When performing the test (code can be found in appendix), there can be observed the following values:

Variable	n	statistic	df	p
People Infected	5762	542.1005	1	6.58e-120

Variable	n	Effect size	magnitude
People Infected	5762	0.09394107	moderate

Since $p < \alpha$ (0.05) and effect size shows a moderate magnitude, we have sufficient evidence to reject the null hypothesis of *wearing a mask has no impact in the amount of people infected by the disease*.

Discussion

“How much does the average amount of waiting increase with social distancing in comparison to without?” can be answered with the hypotheses, to this avail, after the testing is done we can reject the null hypothesis that states that the average waiting time is not affected by the patience level of the attendee. Though it should be mentioned that this difference was moderately significant, it still is enough to reject the null hypothesis. Figure 2 shows that the amount of times waited does indeed increase when the patience in ticks increases per attendee. The median of a patience level of 50 ticks lays significantly higher than the median of the attendees with a patience level of 100 ticks. This then provides an insight into why respecting social distancing at all times is a big challenge for society these days, requiring a lot of patience as waiting times increase for most of the activities that society was not used to waiting for.

The results of the experiments related to the second research question *Will wearing a mask have an impact on the amount of people infected*?was decidingly tested by reviewing the amount of people infected by the virus depending on the percentage of attendees using a mask. As we expected, the usage of masks has an impact in the amount of people infected during the event, as we have enough evidence to reject the null hypothesis: *wearing a mask has no impact in the amount of people infected by the disease.*

Both of these questions, relate to the idea of whether festivals could be a possibility during a pandemic while following guidelines; given the results we believe the impact relies on whether social distancing and mask wearing are practiced to perfection, meaning a mask that is a good quality is worn by every attendee and allowing for a distance of 1.5m or more is fully necessary at all times to minimize the risk of infection, these are variables that are almost impossible to control for event managers in real life events.

Conclusion

On the basis of the previously gained knowledge as described in the introduction and the results of the tests conducted on our model, we can conclude the following.
Looking at the results of the test related to the first research question we can conclude that if one wants to organize a music festival during a pandemic, epidemic or anything similar they need to find a patient crowd. This patience is needed to lower the amount of times attendees have to wait, which is important for attendee satisfaction. The more often attendees have to wait the more frustrated they can get which can result in further accidents later on. Even though the waiting times will be longer, attendees are rewarded with knowing that it does not happen often.
Looking at the results related to the second research question we can not conclude that masks positively influence the spread of the virus. This means that masks are sadly not a reliable solution at a music festival to hinder the spread of a virus.

We hope that in future research more innovative solutions will be tested. Even though we all hope that vaccination will defeat the ongoing pandemic as of 2021 as soon as possible, this disaster can still be an important lesson on public health which can be applied to festivals as well. It is of utmost importance to have proper crowd management and since a pandemic as this one is very likely in the future because of our globalising world, we should gain all the experience we can.

Group effort

When we decided on our division of work, we made sure that we equally distributed the work whilst working on our own expertise. An important factor for us is even though we did some work separately, we always discussed what we did so the other person would understand it and thus could explain the entire project on our own as well. We also made sure that we could always ask each other for help and let the other person take over a task when someone did not know how to fix something. In the end we did around the same amount of work and the division of work per person can be found below.

Sofia Fraile Zendejas

- The code
- Experiments
- Conducting the Experiments
- Results
- Discussion

Jeanine Buurma

- Introduction
- Base of model code
- Model description
- Conclusion
- Discussion

References

Klüpfel, H. (2014). *The Love Paradise Disaster*. Springer International Publishing. DOI 10.1007/978-3-319-02447-9__115

Martella, C., Li, J., Conrado, C., & Vermeeren, A. (2017). On current crowd management practices and the need for increased situation awareness, prediction, and intervention. *Safety Science*, *91*, 381–393. https://doi.org/10.1016/j.ssci.2016.09.006

Pouw, C. A. S., Toschi, F., van Schadewijk, F., & Corbetta, A. (2020). Monitoring physical distancing for crowd management: Real-time trajectory and group analysis. *PLOS ONE*, *15*(10), e0240963. https://doi.org/10.1371/journal.pone.0240963

Sheridan, J., Bryan-Kinns, N., Reeves, S., Marshall, J., & Lane, G. (2011). Graffito: Crowd-based Performative Interaction at Festivals. *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '11*. https://doi.org/10.1145/1979742.1979725

Appendix

Code Waiting Time analysis

```
waiting <- read.table("code (4) WaitingTime-table.csv",header = T , skip =
6, sep = ",")
waiting %>%
  group_by(max.patience) %>%
```

```
  get_summary_stats(waitingTime, type = "median_iqr")
bxp <- ggboxplot(waiting, x = "max.patience", y = "waitingTime",
          color = "max.patience", palette = c("#00AFBB", "#E7B800",
"#CC0066" ),
          ylab = "Waiting Time", xlab = "Patience")
bxp
stat.test <- waiting %>%
  wilcox_test(waitingTime ~ max.patience) %>%
  add_significance()
stat.test

sum <- waiting %>% wilcox_effsize(waitingTime ~ max.patience)
```

Code mask wearing analysis

```
masks2 <- read.table("code (4) masks2-table.csv",header = T , skip = 6, sep
= ",")
set.seed(1234)
masks2 %>% sample_n_by(percent.wearing.masks, size = 1)
masks2 <- masks2 %>%
  reorder_levels(percent.wearing.masks, order = c("0", "50", "100"))
summarymasks2 <- masks2 %>%
  group_by(percent.wearing.masks) %>%
  get_summary_stats(count.people.with..status....infected.., type =
"common")

ggboxplot(masks2, x = "percent.wearing.masks", y =
"count.people.with..status....infected..",
          color = "percent.wearing.masks", palette = c("#00AFBB",
"#E7B800", "#CC0066" ),
          ylab = "People Infected", xlab = "Percentage wearing a mask")

res.kruskal <- masks2 %>%
kruskal_test(count.people.with..status....infected.. ~
percent.wearing.masks)
res.kruskal
effsize <- masks2 %>%
kruskal_effsize(count.people.with..status....infected.. ~
percent.wearing.masks)
```