

An Analysis of the City of Madison's Online Meeting Data

Shri Shruthi Shridhar, Terry Ghanty

1 Introduction

The City of Madison Board, Committee, and Commission (BCC) has transitioned their meetings to a virtual format in response to the COVID-19 pandemic. Ultimately, the virtual format has brought on many changes, some of them being that citizens are now not geographically restricted to attend a meeting, are able to attend and actively participate in more meetings, etc.

The first half of this report (section 2) aims to provide descriptive statistics on the nature of the virtual meetings. Next (section 3), we try to analyze the relationship between different factors of the meetings. The final section includes a preliminary model that predicts the number of participants based on factors such as meeting duration, number of agenda items, starting time, day of the week, and whether the meeting relevant to certain keywords. To provide a better understanding of the characteristics of such meetings, we analyze the two datasets. The first pertains to the characteristics of each meeting (starting time, meeting length, date, committee, etc.), while the other consists of public comments made by participants.

We begin our analysis by considering the following questions:

1. Are there enough metrics to predict the number of participants per meeting? What factors affect number of meeting participants? (Section 2 and 4)
2. How are meeting characteristics different across committees? (Section 2 and 3)
3. From where are the attendants attending this meeting? (Section 2)
4. What are the views of the participants on a high level? (Section 3)
5. Is there a relationship between number of participants and the other meeting metrics? (Section 4)
6. What topics attract more participants? (Section 2 and 4)

2 Meeting Characteristics

2.1 Geography

From Figure 1, we can see the areas which attract the most participants for all meetings across Madison, i.e. overall meeting participants distribution across zip codes. Based on the agenda item and the issue being addressed, we can use the zip codes of the participants to prioritize opportunities in meetings: if we are addressing an issue related to a particular area, we can prioritize letting participants in that area in the meeting room, give them a chance to present their views, etc.

More than 8% of people living in 53703 (Downtown) participate in the meetings and the percent decreases as we move further away from the center of the city. Note that this is under the assumption that the participants in meetings are distinct. In reality, it could be

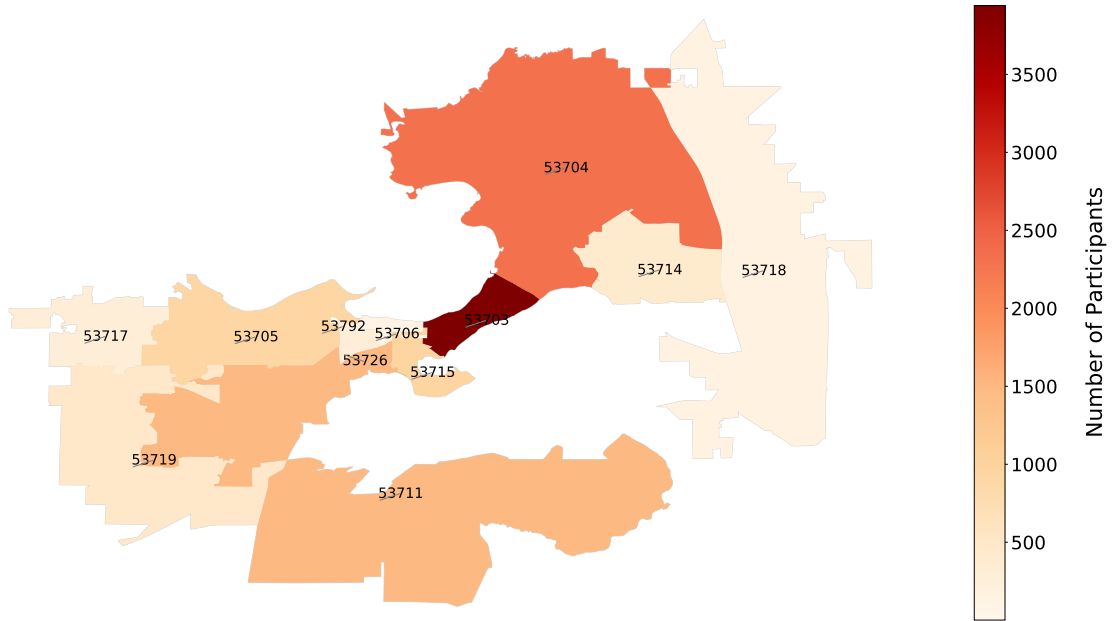


Figure 1: Participants distribution in Madison from all meetings

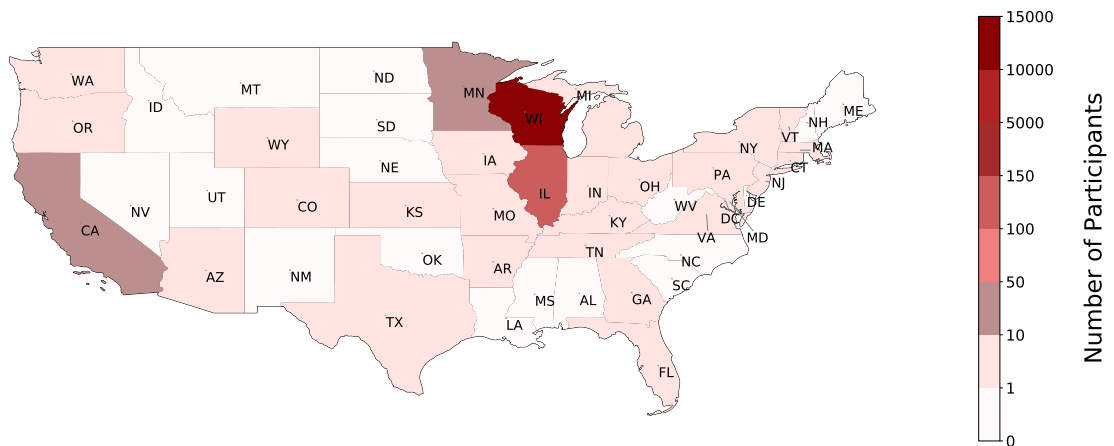


Figure 2: Participants distribution in USA from all meetings

the case that there are individuals who show up regularly. We can try to analyze approaches to attract more people from outside of downtown in the meetings. We plan to further analyze this for committees/boards that attract the most participants.

We also noticed participants from other states and have plotted that in Figure 2. We found that these meetings that attracted other state participants were scheduled to start between 4.30 pm and 6.30 pm and most of these meetings were scheduled on Monday and Wednesday. We can use these dates and times if we are looking to accommodate more participants from other states or avoid these times if the meetings are to be confined to Madison/Wisconsin participants. We analyzed the committees that participants from other states attended and found the average number of participants from other states for these committees (Figure 3).

We analyzed the agenda items that attracted participants from other states and the stand they took which respect to these agenda items they were interested in. We mainly analyzed the top 3 committees that other state participants attend: Urban Design Commission, Plan Commission and Common Council. The agenda items that the other state participants were interested in were related to developments, changes to heating/cooling plants, power pro-

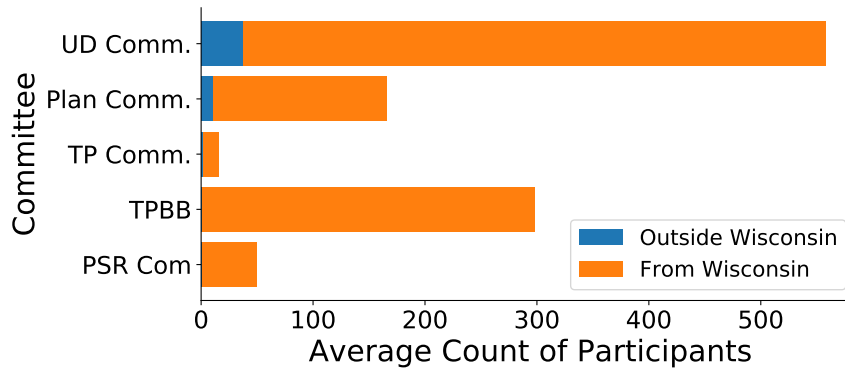


Figure 3: Number of Participants from Within and Outside Wisconsin

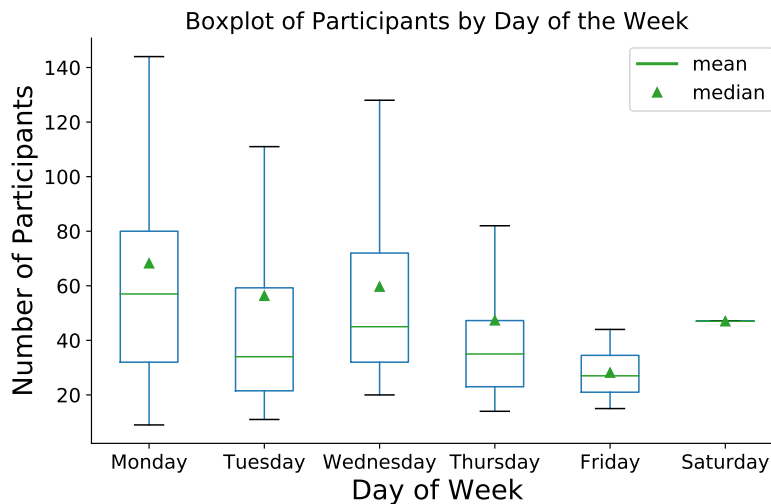


Figure 4: Participation by Day of the Week

duction, mall permits, parks etc and changes related to the police department in the City of Madison. These suggest that the other state participants attending these meetings might either have investments/properties in Madison or temporarily located in outside Wisconsin/have roots in Madison. Their views on the mentioned agenda items (Oppose/Support) mostly matched with the views of those from Wisconsin. This indicates a need for further analysis to understand these participants needs or reasons for participation. We can have separate fields for participants to indicate their state and the reason for their participation if they are from outside Wisconsin. Overall, these analyzes will help us either nurture or avoid participants from other states depending on the committees plan for the interested agenda items.

2.2 Time

Figure 4 gives us an idea of which days of the week would generally render the most number of participants. There seems to be a pattern that as the week progresses, the average number of participants would decrease with the exception of Saturday. This shows us that Monday meetings tend to have the most participants on average. However, there can be other factors at play such as the meeting topic. Additionally, we include the median in addition the mean since the distribution of the number of meeting participants is heavily skewed left meaning that there are a few meetings with an abnormally high number of participants.

Figure 5 shows us the average participation by hour of the day. It is reasonable to divide

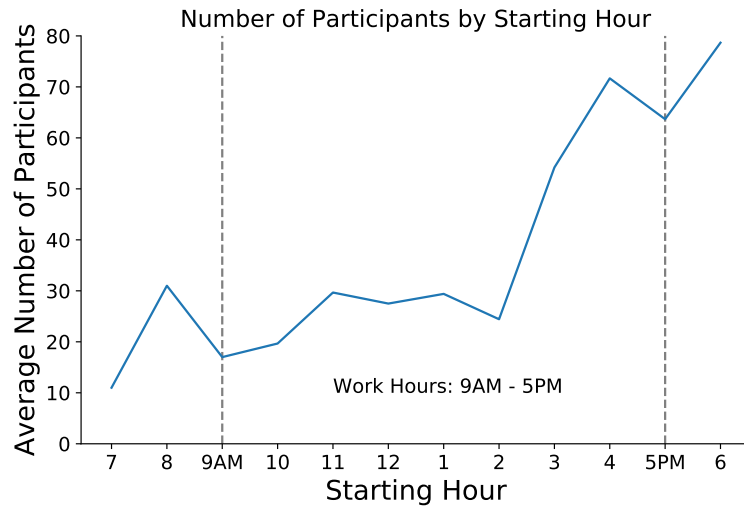


Figure 5: Participation by Hour of the Day

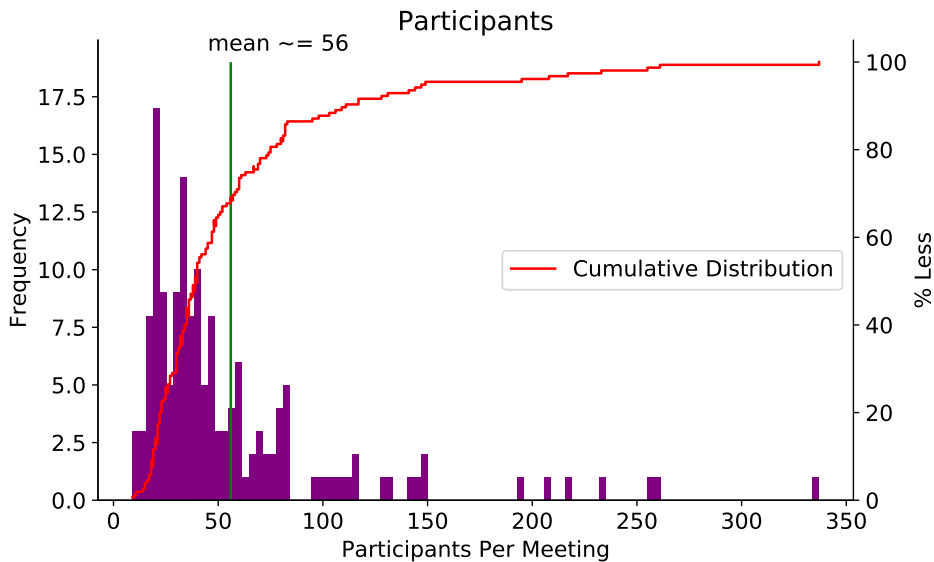


Figure 6: Histogram of Number of Meeting Participants

the day into three categories: before work, during work, and after work. We see that on average, participation is generally lowest around mid-day and increases drastically as the workday ends. An explanation for this trend could be that people are mostly available to attend meetings before and after work.

2.3 Understanding Average Participation and Duration by Committees

For any given meeting, there is an average of 56 participants. Note that the data is heavily skewed to the left meaning that there are certain meetings that have a large number of attendees. These “outlier” meetings are also paired with abnormally long meeting times. The top 10% of all meetings account for 32.02% of all participants.

Figure 7 displays the Average Duration of Meetings in Hours and the Average Participant Counts for each Committee meetings. This helps us identify the committee that draws a lot of participants and the ones that take up more time for further analysis and planning. The Committee name in the plots are sorted in descending order by the participant counts and the

Average Participant Count and Duration during Meetings for each Committee

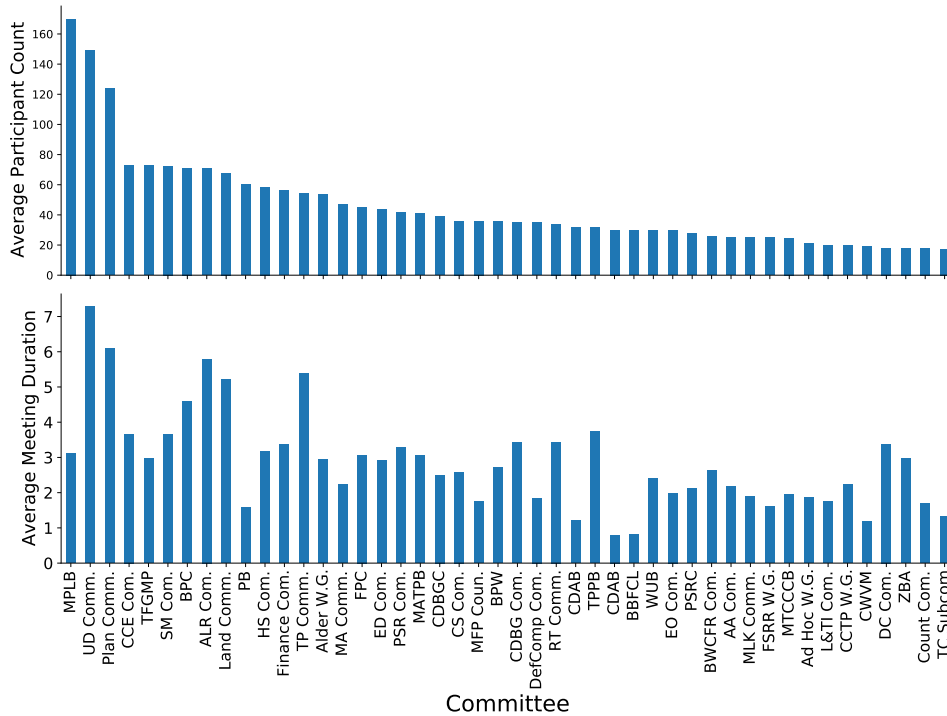


Figure 7: Average Participant Counts and Duration in Hours for each Committee

duration hours. This shows that the highest average participant counts were drawn by the Public Library Board, Urban Design Commission and Plan Commission and the committees that took the highest meeting duration were Urban Design Commission Plan Commission and Transportation Commission. We further analyzed if there is a relation between the average number of participants and the average duration of meetings.

3 Meeting Contents and Further Committee Analysis

3.1 Agenda Items vs. Minutes Items

In the Figure 8, we have plotted the Agenda Items and Minutes Items extracted from the Agenda and Minutes pdf of each item grouped by their committees/boards. The Committee Names are sorted in descending order by the number of agenda items. We see that for most of the committees, the number of agenda items actually addressed during the meeting represented by the minutes items is equal to the agenda items planned. Common Council Executive Committee, Board of Public Works, Plan Commission and TFOGS are not able to address all of their agenda items. There are some outliers where there are no agenda items mentioned and the minutes items are mentioned or when there are more minutes items than agenda items due to addressing of previous or future agenda items in a meeting. These outliers need further analysis. Overall, this plot helps us understand that few committees/boards need restructuring of their meetings to split the agenda items and better refactor their meetings for agenda-items vs. time feasibility.

In the Figures 9 and 10, we can see that most of the datapoints of clustered in the bottom left due to the existence of outliers. We removed the outliers which are defined as points to the left of $Q1 - (1.5 * IQR)$ and to the left of $Q3 + (1.5 * IQR)$ where $Q1$ is the first quartile, $Q3$ is the third quartile and IQR is the Inter Quartile Range defined as $Q3 - Q1$.

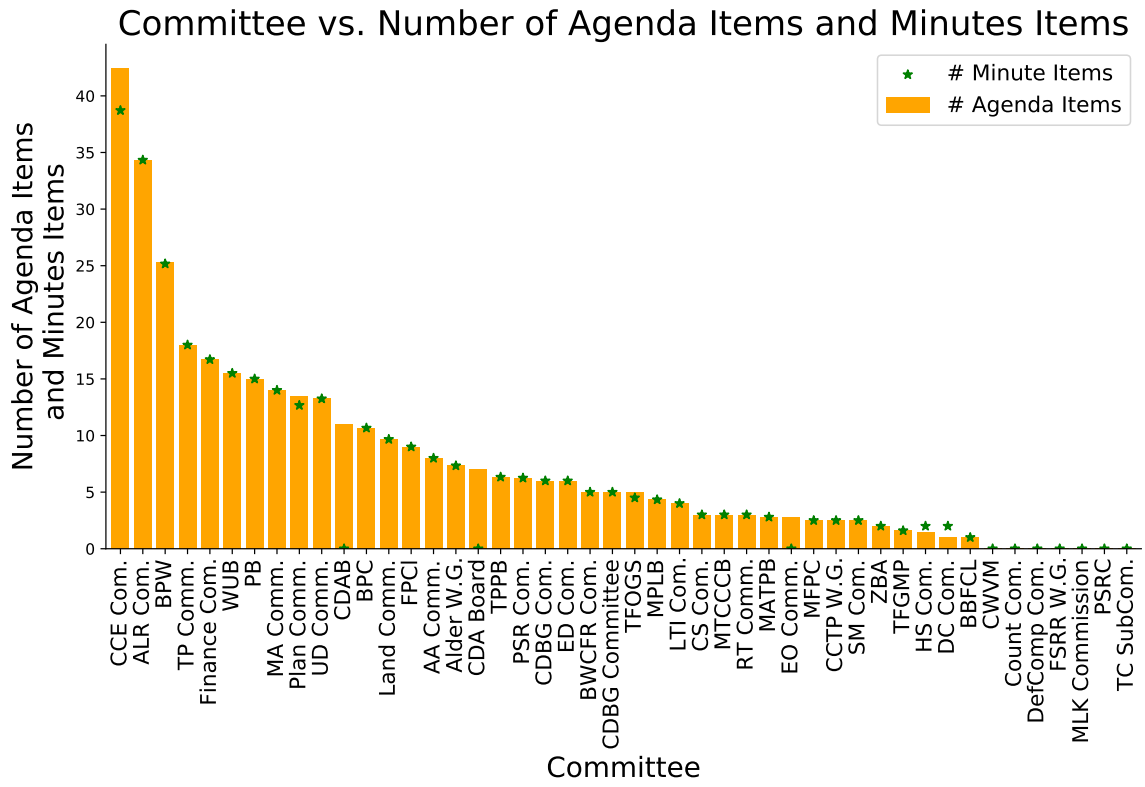


Figure 8: Agenda Items and Minutes Items for each Committee

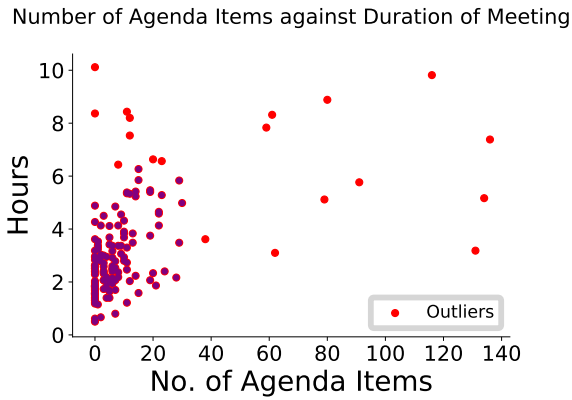


Figure 9: Number of Agenda Items vs. Average Meeting Duration

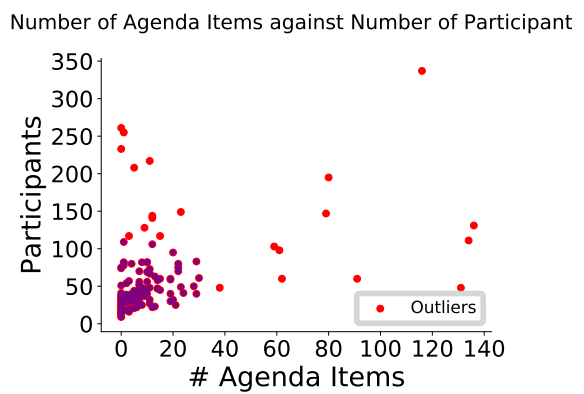


Figure 10: Number of Agenda Items vs. Average Participants Count

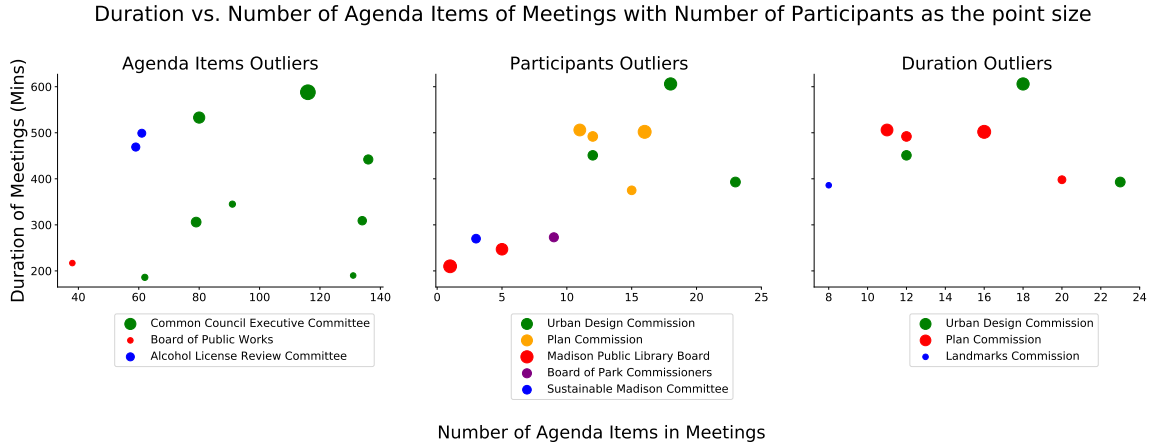


Figure 11: Outliers Characteristics

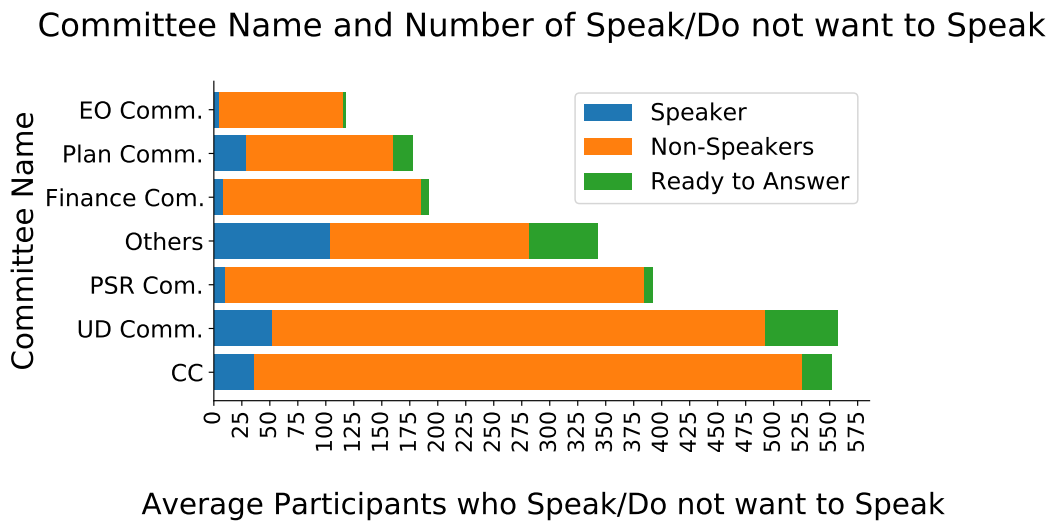


Figure 12: Speaker, Non-speaker counts grouped by Committee/Board

We removed the outliers for participants, hours and agenda items and plotted Agenda Items vs. Participants and Agenda Items vs. Duration below to check for relationship between these features in section 4.

From the agenda outliers subplot in Figure 11, we see that Common Council Executive Committee have frequently long number of agenda items in their meetings and their meetings last more than 3 hours. Considering that the Committee decides guidelines over all the other committees, etc and have regular meetings already, they might not be able to refactor their agenda items much. The other committees that have become an outlier in this regard can considering refactoring their agenda items.

From the participant outliers and duration outliers subplots in Figure 11, we see that Urban Commission Board, Plan Commission and Madison Public Library Board attract a large number of participants and their meetings are also long for a relatively smaller number of agenda items. They can consider refactoring their agenda items across meetings to reduce the duration of the meetings and accommodate priority based participants as discussed in section 2 if there are issues with the online meeting platform.

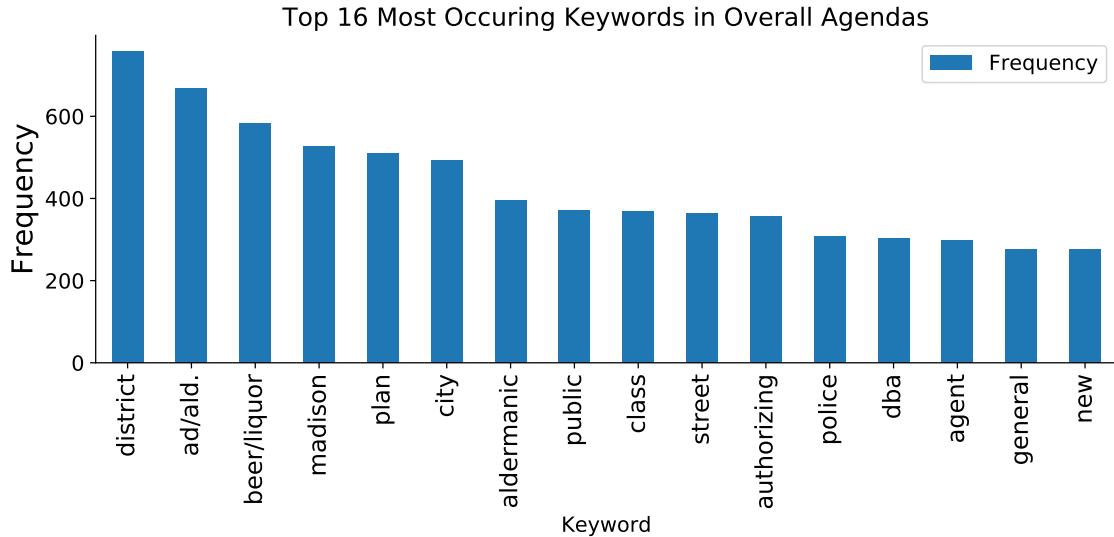


Figure 13: Top 16 Most Frequent Keywords in the Agendas

3.2 Number of Speakers and Non-Speakers

In the above Figure 12, we have plotted the average number of speakers, non-speakers and answering members for each committee/board during all of their meetings. The reason behind this analysis is because in online meetings, the meeting room space is often limited to 500 or 100 people at the maximum. By understanding the number of people who actively participate in the meetings (want to speak or answer), we can create a priority entry based on Speaker/Non-Speaker aspect and other aspects such as first-sign up, etc. to the meeting room when the meeting room is full.

From the plot, we can see that for almost all the committees/boards, especially the ones with higher participation, the number of non-speakers are high. Hence, we can create a priority entry system for the non-speakers when the meeting room is full.

3.3 Number of Supporters and Opposers

We also plotted the average number of supporters - opposers for each committee/board for agenda items during all of their meetings. From the plot, we can see that for some committees/boards (farther left), there are a lot of participants opposing some agenda items (can be analyzed from their registration form). We can use the registration data to find the agenda items garnering a lot of opposition and taking steps to understand their perspective on this item by giving them a chance to speak or voice out their opinion through some forms, online platforms, etc. The agenda items mentioned by these opposing participants also helps us understand the key factors that people care about. The keywords in these agenda items can be used if and when the committee tries to divide their meetings without losing active participation by equally dividing these key agenda items across divided meetings.

3.4 Keywords Analysis/Topics

Figure 13 above shows the top 16 keywords that were mentioned in all of the agenda pdfs. Notice that this list has a lot of noise. Some words are incomplete. However, important words include “district,” “police,” “beer,” “liquor,” and “street.” The next step would be to develop context around these words and use them as inputs in the multiple regression model to predict the number of participants.

4 Forecasting Participation

In this section, we developed a two models in addition to those discussed in Figures 14 and 15. Before we discuss these two models, we attempt to fit a line in the data from Figures 9 and 10 after removing the outliers. The first model presented after that in Figure 16 tries to predict participation based only on one factor: meeting duration. The next model in Figure 17 more factors such as meeting characteristics (section 2), committees, and keywords.

4.1 Understanding Agenda Items and its Relation to Duration and Participation.

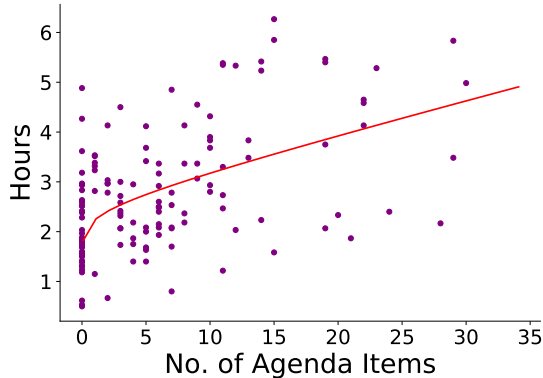


Figure 14: Number of Agenda Items vs. Average Meeting Duration without Outliers

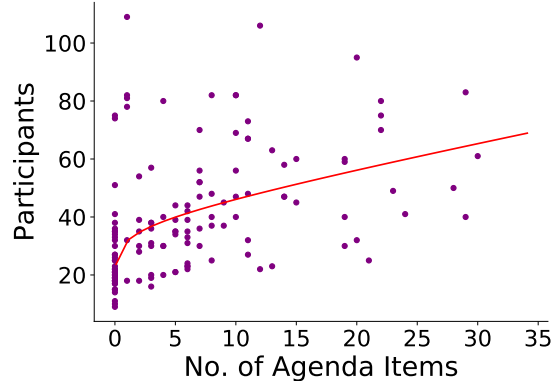


Figure 15: Number of Agenda Items vs. Average Participants Count without Outliers

From the Figures 14 and 15, we see that there is not much relationship between Agenda Items and Participants and Agenda Items and Duration even after outliers removal. The explained variance score in Figure 14 is 0.13005 while in Figure 15 it is 0.1851. Obviously the more agenda items, the meeting will take longer to discuss and there will have a higher chance of attracting more participants. A meeting with 30 agenda items has a predicted number of 65.486 participants. Cutting that down to 15 agenda items renders a predicted number of 51.13 participants. We mention 30 agenda items particularly because the cut off for outliers is 31 agenda items. As we have shown in Figure 11, the Board of Public Works and Alcohol Licence Review Committee can utilize this method of forecasting to reduce meeting duration whilst still maintaining active participation. Also, when looking at the committees with participant outliers in Figure 11, the Urban Design Commission and Plan Commission could also potentially benefit from this.

The line fitting Figure 14 is given by:

$$MeetingDuration = 0.0641 * x + 0.1555 * \log x + 2.17355$$

The line fitting Figure 15 is given by:

$$MeetingParticipants = 0.7871 * x + 3.1594 * \log x + 30.9217$$

where x is the number of agenda items.

4.2 A Deeper Analysis of No. of Participants vs. Meeting Duration

We performed a least-squares polynomial regression in Figure 16 to find a relation between participant count and the duration of the meetings. After multiple trials, we concluded that a second degree polynomial function with Participants in y-axis and Duration in x-axis

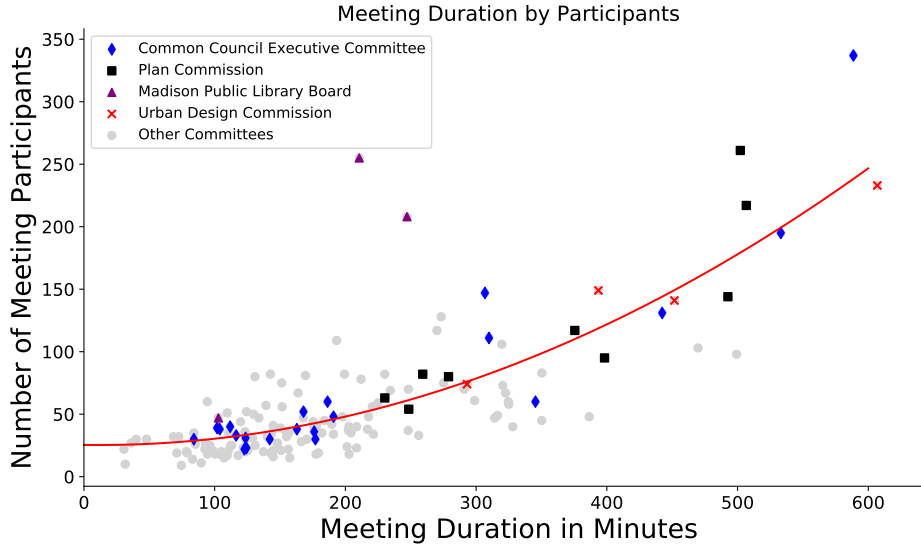


Figure 16: Participants vs. Meeting Duration Colored by Committee

seemed to fit the data best. Each point is representative of a meeting by a committee or a board.

Figure 16 highlights the committees which hold the meetings with the highest number of participants. These meetings are usually paired with longer meeting times as well. Out of the handful that have extremely long meeting duration and high participants is usually hosted by the Common Council Executive Committee, Plan Commission, Madison Public Library Board, and Urban Design Commission.

The equation of the line is the following:

$$MeetingParticipant = -0.0143971116 * x + 0.0006387333 * x^2 + 25.4139195339048$$

where x is the Duration of the meeting in minutes.

Figure 16 gives us a general relationship between the meeting duration and the number of meeting participants. The line of fit is created using a second degree polynomial function. This creates a better fit than a linear trend. What this tells us is that as the meeting duration increase, the model shows an approximate exponential increase in the number of participants. Notice that meetings beyond around 250 minutes tend to be more scattered than meetings less than 250 minutes. The explained variance for this model is 0.61156. We attempt a more complicated model to improve this.

4.3 A Lasso Regression Model to Predict Number of Meeting Participants

The above shows a simple model in which the predicted value of meeting participants is depends only on meeting duration. This is meant to illustrate a positive but squared-increasing relationship. Generally, the above model states that the longer the meeting duration, the more participant. However, this is not enough to predict participation since one you know the meeting duration, you will know the number of participants.

Our goal is to create a model that can accurately predict the number of meeting participants based on some key factors: meeting characteristics, keywords/meeting topics, and the type of committee. Our first attempt was a multiple linear regression where each feature that is used is a factor in the model. The problem is that there were too many features and

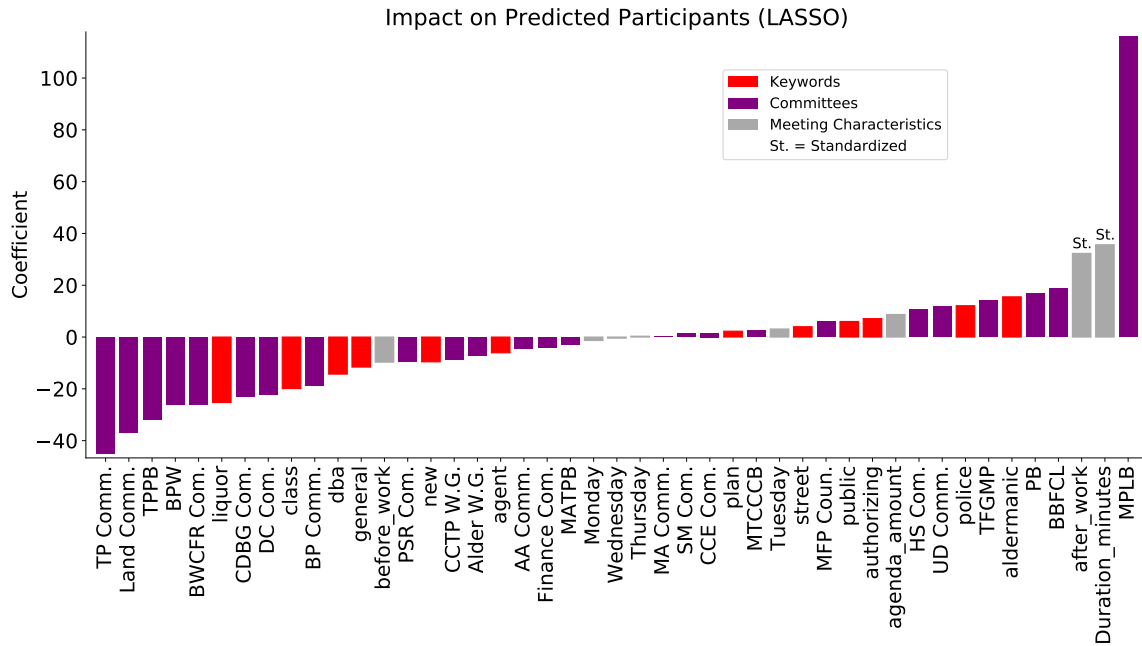


Figure 17: Coefficients of the LASSO Regression Model to Predict the Number of Participants

it created a risk of over fitting the data. Thus below is a plot of the different features of the LASSO regression. The LASSO regression is similar to the multiple regression but it removed 27 trivial factors.

The model of the LASSO Regression has the following equation:

$$\begin{aligned} \text{Number of Participants} = & 8.715 * (\text{agenda amount}) + 35.630 * (\text{Duration minutes}) - 9.697 * (\text{before work}) \\ & + 32.301 * (\text{after work}) - 1.333 * (\text{Monday}) + 3.112 * (\text{Tuesday}) - 0.431 * (\text{Wednesday}) \\ & + 0.330 * (\text{Thursday}) + 5.828 * (\text{MFP Coun.}) + 18.778 * (\text{BBFCL}) + 1.289 * (\text{SM Com.}) - \\ & 7.160 * (\text{Alder W.G.}) - 2.927 * (\text{MATPB}) + 2.465 * (\text{MTCCCB}) - 36.921 * (\text{Land Comm.}) - \\ & 18.824 * (\text{BP Comm.}) - 4.300 * (\text{AA Comm.}) - 31.773 * (\text{TPPB}) - 8.612 * (\text{CCTP W.G.}) + \\ & 16.663 * (\text{PB}) - 25.904 * (\text{BPW}) + 10.808 * (\text{HS Com.}) - 23.138 * (\text{CDBG Com.}) + 116.206 * (\text{MPLB}) \\ & + 14.191 * (\text{TFGMP}) + 1.568 * (\text{CCE Com.}) - 22.268 * (\text{DC Com.}) + 0.367 * (\text{MA Comm.}) - \\ & 45.049 * (\text{TP Comm.}) - 25.880 * (\text{BWCFR Com.}) - 4.168 * (\text{Finance Com.}) + 11.842 * (\text{UD Comm.}) \\ & - 9.616 * (\text{PSR Com.}) - 25.293 * (\text{liquor}) + 2.260 * (\text{plan}) + 15.527 * (\text{aldermanic}) + \\ & 6.016 * (\text{public}) - 19.898 * (\text{class}) + 3.992 * (\text{street}) + 7.119 * (\text{authorizing}) + 12.100 * (\text{police}) + \\ & 14.313 * (\text{dba}) - 6.096 * (\text{agent}) - 11.592 * (\text{general}) - 9.571 * (\text{new}) + 55.080 \end{aligned}$$

This equation is visualized in Figure 17. Note that:

- Two columns, agenda_amount and Duration_minutes were standardized. So an increase in 1 minute doesn't mean a predicted increase in 35.6298 participants.
- The LASSO sent the 27 coefficients of during_work, Friday, Saturday, ALR Com., RT Comm., CWVM, ZBA, FPC, DefComp Com., EO Comm., CDBGC, Count Com., WUB, FSRR W.G., L&TI Com., PSRC, TC Subcom., CS Com., Ad Hoc W.G., ED Com., CDAB, MLK Comm., district, alder, madison, city, and Plan Comm. to 0. This indicates that these factors do not influence predicted participation.
- (Agenda Amount) is the number of agenda item planned for that meeting.
- The topics/keywords (outlined in red in the Figure) are encoded as 1 if the word is mentioned in the meeting. These are the top words as sorted through the Agenda

Items, applied onto the Minute Items. Additionally, characteristics such as starting time of the meeting and day of the week are also one-hot encoded.

- The committees follow the same one-hot encoding where it is 1 if the meeting is hosted by that committee.
- The explained variance is 0.79911. 0.18755 higher than the simpler model in Figure 16.

We see that the Madison Public Library Board has the most impact on the predicted number of participants. Additionally starting time has a positive impact on the predicted number of participants as well. This could indicate that meetings that hope to have a higher number of participants should host their meetings towards the end of the workday. The model predicts that meetings after work can have an increase in about 32 participants. Words such as “aldermanic”, “police,” and “authorizing” also have a positive impact. Additionally, the model shows that Tuesdays and after_work have the highest impact on participants with an expected increase of 3.112198 and 32.300564, respectively. Furthermore, if a meeting is discussing something related to an aldermanic district and police, the model predicts an increase in 15.526517 and 12.099811 participants, respectively. We hypothesize that people are interested in meetings that are dealing with local issues since aldermanic is tied to aldermanic district.

5 Conclusions and Recommendations

From the descriptive statistics and analysis of the two datasets, we can see that:

- We see more participants for meetings when they start after work hours and towards the start of the week. For meetings that cover important issues and would benefit from more participation, schedule meetings towards the end of the work day (4PM onwards) and earlier on in the week (Monday tends to have a higher average participation).
- More people attend meetings addressing local issues. The LASSO model points to keywords such as police can increase meeting participation by about 12 people. Additionally, issues concerning an aldermanic could increase participation by about 16 people.
- We see that some committees/boards such as Common Council, Urban Design Commission, Plan Commission, Madison Public Library Board have longer meetings than the others. They can consider refactoring their meetings and meeting more frequently to reduce the duration of their meetings. According to the fit in 15 we see that though long meetings have more participants, meetings with less agenda items can still maintain a high number of participants. A meeting with 30 agenda items has a predicted number of 65.486 participants. Cutting that down to 15 agenda items renders a predicted number of 51.13 participants. Thus, the committee can have two meetings to cut down the meeting time and still have more participants in total.
- More than 8% of people living in 53703 (Downtown) participate in the meetings and the percent decreases as we move further away from the center of the city. Meetings should be further promoted in these areas.
- We have people from outside of Madison and other states attending the meetings as well, for 20 committees. According to the data, meetings that attract the most out of state participants start on Wednesdays around 4:30 PM. This can help us accommodate meetings attracting them separately or at a convenient time that works across time zones or prevent other states participants giving preference to local ones, based on

committee preferences. We see that the other state participants' views are similar to Wisconsin participants and these participants are mainly interested in agenda items relating to developments/changes to Madison based properties/departments. Based on this, we recommend having a separate field for participants to mention the reason for their participation if they are from outside Wisconsin.

- There are more non speakers than speakers for most of the meetings. In case of an participants overflow, the speakers can be given priority entry to the meeting to allow for an active discussion. We recommend having a separate field in the registration form for participants to reason out their stand regarding the agenda item they are interested in accommodating the possibility that non speakers are not interested in publicly voicing their opinions/reasons.
- For each committee, we can understand the overall stand taken by all the participants (majority) for each committee. This can be further analyzed by finding the stand of participants for each meeting and even, each agenda item. This can be helpful to understand the participants view and opinions on a high level.

Appendix A Acronyms

CWVM: City-Wide Violence Meeting
WUB: Water Utility Board
C Com.: Complete Count Committee
MPLB: Madison Public Library Board
DC Com.: Downtown Coordinating Committee
MFP Coun.: Madison Food Policy Council
RT Comm.: Room Tax Commission
L&TI Com.: Landlord and Tenant Issues Committee
PB: Personnel Board
CDAB: CDA Board
TC Subcom.: Traffic Calming Subcommittee
AA Comm.: Affirmative Action Commission
Ad Hoc W.G.: Ad Hoc Task Force on the Structure of City Government (TFOGS) Final Report Implementation Work Group
CDAB: Community Development Authority Board
CDBG Com.: CDBG Committee
BPW: Board of Public Works
BP Comm.: Board of Park Commissioners
PSRC: PSRC
ALR Com.: Alcohol License Review Committee
BWCFR Com.: Body-Worn Camera Feasibility Review Committee
Finance Com.: Finance Committee
CS Com.: Community Services Committee
MATPB: MATPB
Plan Comm.: Plan Commission
ZBA: Zoning Board of Appeals
Alder W.G.: Alder Workgroup to Develop Logistics and Operational Details for MPD Independent Civilian Oversight
Land Comm.: Landmarks Commission
PSR Com.: Public Safety Review Committee
CCTP W.G.: President's Work Group on Council Communication Tools Processes
FPC: Food Policy Council
TPPB: Transportation Policy and Planning Board
CDBGC: Community Development Block Grant Committee
ED Com.: Economic Development Committee
TFGMP: Task Force on Golf in Madison Parks
TP Comm.: Transportation Commission
UD Comm.: Urban Design Commission
BBFCL: Board of Building Code, Fire Code, Conveyance Code and Licensing Appeals
MTCCCB: Monona Terrace Community Convention Center Board
HS Com.: Housing Strategy Committee
DefComp Com.: Deferred Compensation Committee
MLK Comm.: MLK Commission
CCE Com.: Common Council Executive Committee
SM Com.: Sustainable Madison Committee
MA Comm.: Madison Arts Commission
EO Comm.: Equal Opportunities Commission
FSRR W.G.: Food Systems Recovery and Resilience Work Group