ABSTRACT
The Toronto International Film Festival (“TIFF”) is an annual event, screening a huge variety of new films for the international film industry as well as the general public. The number of choices means selecting which films to order tickets for can be overwhelming. I suffer the occupational hazard of considering every logic problem in terms of SAS® code. Here we explore how to use some very simple code to explore scheduling options, which will support decision–making with the goal of seeing the most films from a priority list in the most enjoyable way. While many of us use SAS for efficiency in our work, this is a small example of how it can also be beneficial for personal time.

BACKGROUND
TIFF has celebrated its 37th annual festival. It keeps growing to include more films at more venues over 11 consecutive days each September. In 2012, there were multiple screenings of each of over 300 distinct films from more than 60 countries. That can make scheduling and ticket selection for a casual filmgoer overwhelming. Most venues screen 4 or 5 events per day. Films are scheduled daily beginning between 8:30 or 9:00 AM and continuing through to a Midnight Madness track. A sample from the festival program is shown on page 3. It looks like a random mosaic -- and was about as logical to work with as one. Each day covers 2 pages like that. So, with luck, good timing and the sacrifice of comforts like regular meals, sleep and exposure to daylight, it would be possible to attend up to 6 screenings per day. That reminds me of what I tried to do at my first SUGI conference, where I had every hour of the day packed with presentations and it was a bit of a blur by the end. There’s a practical limit to how many screenings or presentations a person can fully enjoy in a day.

Tickets for festival screenings may be purchased individually, but discounted blocks of passes can be ordered for various packages (unlimited day and evening use, daytime screenings use only, second half of the festival only) based on the expected popularity of those packages. Typically, the cost–per–pass declines with the number of passes in the package as well as the popularity of the package.

The ticket acquisition process is a complex one. TIFF systems have evolved over the years to provide lots of information and much of the selection process online. Still, there are many steps and long waits to order tickets:

• Decide who will participate in sharing a block of passes. Passes or tickets are transferrable and can be shared with others, but for the sake of logistics, consider a practical maximum of 2 – 4 primary participants in a group to share a block of passes.

• Decide which type of package and how many passes to order.

• Order the block of passes.

• Receive confirmation of the order and information on some of the next steps.

• Wait (weeks)! While waiting, prepare for ticket ordering by thinking about which films are priorities and how the decisions will be made. Make preliminary decisions and have alternate selections ready.
• Receive notification of when the online ticket selection window assigned to the order begins. To manage traffic on the website or at the call centre, each order gets a staggered earliest allowed start time to begin the actual ticket ordering. These windows are assigned randomly within priority group. That is, orders for the most expensive packages get a random start time within their window first, and then orders for less expensive packages get random start times within subsequent windows. Once logged in to the ticketing system, a user has a maximum of one hour to complete all selections. What stress – only one hour to navigate the system and make all selections! This is why being prepared is important. Any passes not converted to tickets by the end of the hour can be exchanged for tickets remaining in inventory after all users have completed online ticket selection.

• Wait again (days) for the assigned ticket selection window.
• When the appointed hour arrives, log in and select tickets.
• Receive email confirmation of tickets ordered; agonize over the decisions made and consider trying to exchange some of the tickets later.
• Wait some more.
• During the days immediately before the festival begins, wait in line to pick up the set of printed tickets ordered. Of course, standing in line to pick up tickets with lots of other people who are doing the same thing leads to friendly discussions about who has chosen which films and why. There’s even enough time during the line-up to reconsider the choices and check whether there are still tickets available to something else.
• Optional extra: Wait in another line to exchange some of the printed tickets for a different film or make use of vouchers not assigned during the ticket selection hour.

To make the most of the festival, lots of planning and preparation was required days and weeks before the lights were dimmed for the first screening.

THE CHALLENGE

Different people might have very different objectives for their festival experience: Industry insiders would come with a very specific list of priorities. Members of the public, depending on how serious they are about film, might be hoping to maximize the number of films viewed (either in total on or a particular day), follow a themed track, see the most stars and famous people, attend the most Q&A sessions, or enjoy a relaxed and casual adventure.
Many festival-goers invest lots of time poring over the film catalogue to identify their preferences and then go through the ordering process with no other preparation, but with factors including what else is screening at the same time, how many times a film will be screened during the festival, what time of day it’s offered, where it’s offered, and what our personal schedules allowed to consider, we decided a systems approach was in order.

In order to share a block of film passes, we needed to agree on how many passes and decide how to decide on which films to see. We agreed that each person would plan to see 10 films during the festival. It would have been possible to split the block of passes so that not everyone attended the same screenings, but the consensus was to go along with the group. That certainly simplified the order, but still, the plan needed to be fair – and a plan was definitely required.

After surveying each of the primary participants for their “Top 20” preferences for which 10 films to see (from a list of over 300), inevitably someone will come back with a list of only about 5 films they really care about and agree to go along with the rest to whatever gets selected. Chances are that someone will come back with a “Top 20” list of 21 names, too. People are human, after all. We dealt with this by assigning a fixed number of preference points to each primary participant. Each of their preferences was weighted depending on the total number of choices they flagged. For example, if participants were asked to choose 20 films and chose 20, each of their preferences would be assigned a weighting of 5 points (100 points/20 = 5). If one person only chose 4 preferred films and agreed to accept the preferences of others for the rest, each of their choices would be assigned a weighting of 25 points (100 points/4 = 25).
We also needed to determine our desired interval between start times of films. Did it make sense to see one film at 9:00 AM and another at 11:30 AM? Although the approximate length of each film and the location of each screening were available on the website, it wasn’t structured in a way that was easy to collect. Also, it was impossible to anticipate from the data whether there would be a Q&A session following a film. Whether there would be a Q&A follow-up and how interesting it might be depended on whether the writer(s), producer(s), director or actor(s) happened to be available. Some came to Toronto but stayed only for the premiere of their film; others stuck around to enjoy the whole event with its associated parties. Travel time between venues, allowing time to line-up for seats at the next theatre, and practical breaks for meals or snacks were also considerations. We worked with a broad assumption, which dictated that we’d allow at least 4 hours between start times of screenings we planned to attend. Most of the films were more than 1 hour and less than 2 hours in length, but each screening began with brief comments from the programmer and possibly an introduction from the writer, producer, director or star (if they were present) and a follow-up Q&A session with them. All of the venues are in downtown Toronto and most are within easy walking distance of one another. The second screening we chose each day might be in the same building as the first or might be 30 minutes away. A ticket was for admission to the venue but not for an assigned seat, so line-ups of ticket-holders formed outside the theatre well in advance of the screening time. Some of these line-ups began more than an hour before the screening time, but this was difficult to anticipate. While it might have been possible to see a film, leave right after the screening and get into the next theatre in much less time, we agreed that a 4 hour window would be comfortable and would even allow for meal or snack breaks.

Now, with an amazing selection of films, the desire to use ten passes per person over the 11 days, some rules about how to schedule the days – and the occupational hazard of seeing every logic problem in terms of SAS code - we looked at how SAS could help make ticket selection and ordering easier. Unlike many situations when we use SAS for decision support or problem solving, there is no one correct solution to this challenge and no one way to measure a successful outcome. SAS was used to generate possible scenarios for review rather than identify one solution, since there are so many soft factors involved in optimizing enjoyment. The need to gather preferences and attach them to possible screenings immediately made me think of joining tables in SQL. Perhaps that’s another occupational hazard.

**TACKLING THE DATA**

Good data support good decisions, so we set out to organize some usable data. Enterprise Guide® software (“EG”) was used to explore and organize the data on the fly and generate result sets from the preferences of the participants. The program nodes in EG were combinations of DATA steps and PROC SQL. Note that all of this work with the data falls within the ‘prepare for ticket ordering’ step described above as part of the ticket acquisition process. Since most festival-goers would not have approached ordering tickets as a data management challenge, the data were not organized in an ideal way for automated processing. There was, however, a chronological list of films for each day of the festival on a separate web page.

Our process:

1. Copy and paste the daily schedule tables from the website into one concatenated list of all available films and their screening dates and times.

2. Import the concatenated list into SAS.
3. Filter the data: Are there times the group members aren’t available or don’t wish to attend? Drop rows for these times. Call this data set “FILMS”.

4. Add a counter, TIMES_SCREENED, to each row of the FILMS table to indicate how many showings of that film remain. (A film with the same number of PREFERENCE_POINTS as another but a smaller number of TIMES_SCREENED would be sorted into priority sequence to increase the likelihood of successfully scheduling it.)

5. Clean up the data: Based on an initial attempt, it was clearly necessary to do some data cleaning. Some of the film titles were wrapped in quotes, some had leading blanks, and there was inconsistent use of “The” as part of some titles and inconsistency as to whether a title beginning with “The” would appear under “T” or under the first letter of the next word in the title in the alphabetical descriptions of the films. SAS functions came to the rescue to remove leading blanks and quotes from around some film titles and remove the use of “The” at the beginning of film names for sorting and discussion amongst the attendees (while keeping the original name available as the key because it matched what would be in the order system). Call this SCREENINGS.

6. Identify the distinct set of films from SCREENINGS and sort the list by film name. Save this as “FILM_NAMES”.

7. Export FILM_NAMES to Excel and distribute a copy of the Excel file to each of the primary participants in the group so that they can identify preferences. Ask each participant to populate a PREFERENCE_FLAG column to indicate the film is a priority.

8. Meanwhile, decide on preferred frequency or interval and populate columns called SESSION and FESTIVAL_DAY in FILMS:
   • Based on our agreed assumption regarding how often we’d see films, use the date and time fields to derive SESSION, where each morning, afternoon and evening of the 11 days had its own SESSION value (1, 2 or 3). That ensured we wouldn’t order tickets for two films in the same half day. This was based on a somewhat arbitrary rule that films starting after 8:30 AM and before 1:00 PM would be our ‘morning’, 1:00 PM to before 6:00 PM would our ‘afternoon’, 6:00 PM to 10:00 PM would be our ‘evening’ and we wouldn’t consider any start times later than 10:00 PM. No Midnight Madness fans in this group!
   • Use SCREENING_DATE and the date of the start of the festival to derive and populate FESTIVAL_DAY, which will have a value of 1 through 11 for the first through eleventh day of screenings.

SESSION and FESTIVAL_DAY will provide index values for an array representing each half-day of the festival. This will help keep track of whether we have selected a film for a particular half-day or whether that time is still available.

9. Gather the returned FILM_NAMES spreadsheet from each participant. For each spreadsheet, import into a SAS data set, assign a number of PREFERENCE_POINTS to each row of data to give a
weighting based on the how many (or how few) preferred films this participant chose. Keep only the rows representing films flagged as preferences by that participant.

10. Concatenate the preference data sets from all participants. Call this PREFERENCES1.

11. Summarize PREFERENCES1 by FILM_NAME, with a sum of all the PREFERENCE_POINTS and a frequency count to track how many of the participants identified this film as a priority. Call this PREFERENCES2.

12. Reality check: Examine the summary data here to make sure the participants have reasonably common preferences. If all the films have a frequency count of one or two, it’s time to reconsider whether the participants will be happy seeing all the same films. If necessary, negotiate priorities or reconsider who the primary participants will be and loop back to step 9 until there is reasonable commonality.

13. Join PREFERENCES2 with FILMS on the common key of FILM_NAME. Keep only the rows common to both tables. Call this “POSSIBILITIES”. This will be the list of all screening dates and times for all films the group considers a priority.

Stopping here, with a clean list of possibilities to satisfy the group, would have provided many advantages already. Still, it was easy enough to go further with a SAS array.

14. After all our data preparation steps above, the code shown below does the really powerful processing: The POSSIBILITIES data set, sorted by descending PREFERENCE_POINTS and by TIMES_SCREENED, was used to populate an array containing a row for each day of the festival and a column for each SESSION (morning/afternoon/evening). For each row of the data set, if the data element at SCHEDULE[FESTIVAL_DAY, SESSION] was blank, FILM_NAME would be stored there. Once FILM_NAME is assigned to the array, subsequent rows for the same film will be ignored.

```sas
data SCHEDULE(keep=schedule);   
array schedule[11,3] $80;        
retain;                      
set POSSIBILITIES end=last;  
do i = 1 to 11;         
    do j = 1 to 3;      
      ** Determine whether a different screening of this **;  
      ** film has already been scheduled. If it has, then **;  
      ** ignore this occurrence. **;  
      if film = schedule[i, j] then do;  
        return;  
      end;  
      end;  
    end;  
    if schedule[FESTIVAL_DAY, SESSION] = " " and ticket_count < 10 then do;  
      ** select this screening **;  
      schedule[FESTIVAL_DAY, SESSION] = film;  
      ticket_count +1;  
    end;  
    if last;  
    output;  
run;
```
15. We ran the step above repeatedly, removing the row representing the first occurrence of successful assignment in order to generate other possible schedules for consideration. This was a crude ad-hoc approach but it generated many alternatives we would not have recognized otherwise.

What else might be added? (Proposed future enhancements based on initial experience):

16. A macro to facilitate eliminating data rows for Off Sale screenings: This would still require manual input during the ordering hour but would be cleaner than directly editing the data.

17. A more sophisticated algorithm for selecting which screening for films with multiple screenings during available times could be tried. We avoided generating all possible combinations of screenings of all priority list films because of the volume, but there may be some merit in generating all the possibilities if some kind of automated score can be applied to rank the results. In retrospect, after observing that screenings on the first weekend were the most popular ones, working from latest to earliest screening per film might make more sense than working from earliest to latest.

IMPLEMENTING THE PLAN

After various result scenarios were generated in SAS, they were subject to manual review with judgement applied. Optimal enjoyment is highly subjective and it can be influenced by many factors that can't be predicted as well as some that can be, so the programmatic solution to identify scenarios was only one component of the final selection process. Being prepared with ranked and sorted lists, but also having enough information (meaning organized data) easily accessible to make informed last-minute decisions was very helpful.

The day before our randomly assigned hour to begin ordering tickets, TIFF was trending on Twitter. There was lots of excitement about the festival -- along with comments about some of the popular films that were Off Sale already. This didn't mean the screening was necessarily Sold Out, but no additional tickets were available to order at the moment; returns or Rush tickets might be available once the festival began. Given the drama of selecting screenings so far, we chose not to hold onto any vouchers and risk racing around on the final days of the festival to find screenings that had Rush tickets and line-ups that weren't too long. So, we needed the flexibility to make some quick last-minute decisions during the ticket ordering hour based on real-time information about which screenings were still available.

Finally, the magic hour arrived to place the ticket order. It was a relatively late window for ordering and many of the premier screenings were not available. Still, armed with carefully organized data and the ability to exclude the unavailable screenings and generate alternate scenarios within seconds, an hour was more than enough time to consider alternatives and find subsequent screenings of our priority film choices. We decided that attending two screenings per day with some days off, for example, would be more convenient than screenings on 10 different days. Nine of the ten films ultimately selected were from our priority ranked list. The tenth was chosen because it fit very conveniently into the schedule. After all, writers, producers, directors, actors and others had invested their time and money into the film and a TIFF programmer considered it worthy, so the tenth seemed worth the risk for the convenience. If the choice is between seeing two of top priority films but having to rush from one venue to the other in traffic and
possibly sit in the least desirable seats for the second film because you arrived close to the start time and seeing a top priority film and a slightly lower priority film which happens to be in the same venue with enough time to go for a drink and re-join the line in time to get optimal seats for the second screening, perhaps the lower priority becomes more appealing.

ALTERNATIVE APPROACHES

For those who are completely overwhelmed by doing the research or selecting from a huge number of choices, there are some to limit or eliminate the need for personal ticket selection by following others’ lead:

- Check out the Programmers. TIFF programmers have nothing to do with SAS, but spend their work days in front of screens for a different reason: their job is to view hundreds of new films from around the world each year to identify those which might be appropriate for TIFF. Each of these programmers has a theme for the festival and some cover specific geographic regions. Focusing on the selections from one particular trusted programmer would reduce the number of possibilities to be considered and might heighten the festival experience, but this would require (1) some knowledge of the work of that programmer and respect of their judgement, and (2) gathering data about which programmer was responsible for selecting which film. The data are available for each film but are not in a format that would allow for automated collection.

- Stick with one programme: Programme tracks might be things like documentaries, collections of short films, first-time directors, or world cinema. For someone with specific interests, this would reduce the number of choices, but still require some decisions because one programme might have several simultaneous screenings.

- Stick with one venue: The festival took place on 25 screens at 10 venues. Most of the screens were used 4 or 5 times per day for different films. Although it would be impractical or even impossible to plan to see consecutive screenings in the same venue, sticking with one venue and attending alternating screenings there would simplify the logistics.

- Crowd-source it: Wait until late in the ordering process and keep an eye on which films have high demand for tickets, or follow festival veterans or professional film critics on social media to learn what they’re excited about. Twitter was a great source for information in the days leading up to and during the festival there was helpful information about how long the line-ups outside theatres were, links to reviews and updates about which stars had been spotted in town and were likely to appear at the screening of their film.

- Purchase a Programmer’s Choice package: With these packages, individual film selection is taken out of the hands of the participants. Tickets are assigned based on a selected programme.
LESSONS LEARNED

• It’s helpful to have a flexible schedule and an open mind. Some of the most meaningful moments were the one’s we hadn’t anticipated: Sticking around for an informal Q&A session with the writers and producers to hear them explain their thought process. Because some of these screenings were world premiere, it was especially poignant to hear feedback from the audience and how one of the directors had an epiphany during the Q&A when he tried to explain the background for a character and suddenly realized that the scene portraying that background had been cut.

• Be prepared during the ticket selection process AND during the festival: Carry a bottle of water and some non-perishable food and dress for possible weather changes outdoors during line-ups and for temperature variations inside theatres.

• We found that seeing two films per day was quite comfortable and allowed us to actually remember everything we saw.

• Although most people who order blocks of festival tickets are using scribbled notes as their decision support, there is much to be gained with a bit of technology. Using Excel to manage the contents of the tables would have been a good step; it’s one that is available to people who don’t have SAS, but the ease of joining tables, selecting distinct rows and being able update and store results scenarios very quickly made SAS an advantage.

CONCLUSIONS

Being prepared at all stages from ordering passes through attending the films made the whole process less daunting and more enjoyable. There are lots of other ways to select tickets and have a thoroughly enjoyable experience, but running some very simple SAS code to generate and report on different schedule combinations was a great confidence-builder for the intense decision-making process required to order tickets. An automated approach also supported a fair selection to satisfy all participants in the group. Repeated waiting and line-ups mean the festival is only for the patient, but overall it was a very positive experience.

CONTACT INFORMATION

Your comments are welcome.

Lisa Eckler
lisa.eckler@sympatico.ca

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.® indicates USA registration. Other brand and product names are registered trademarks or trademarks of their respective companies.