Requirements specification for Performance Rating business process: Hilarity Haven Circus

1. General description of business process

a. A general description of the business process:

The process of performance rating is as follows: the customer buys the ticket for a desired performance either at the booth stand or online and attends that performance.
Then they go to our circus website, click to take part in the survey, after that they scan the ticked code or enter the number to the circus survey. It is also possible to get the survey from the booth in the paper form. After validation of the existence of such a ticket, the customer is taken to the survey, in which he answers all or at least some of the presented questions. The data gets collected to the system, and then can be analyzed to enhance the business performance.

The customer can enter the survey for one ticket only once and only after the performance. If they have more tickets, then they can take part in it multiple times.

b. The main business goals of the business process:

1. Rising trend of **the number of people** doing surveys at a level not lower than 0.5% per month.
2. Rising trend for both ratings (**average cleanliness rate and average satisfaction rate**) at 0.3 per month.

c. Typical questions

Compare the overall rating of the customer to the cleanness rating the customer gave.
How many customers were willing to rate each type of performance.
Give the best rated performances for this week.
Give the performer who took part in most of the low rated performances.
Compare type of performances in terms of rating.
Give the average rating of performances in a week.
Depending on type of ticket (reduced, normal or VIP) is there any difference for the rating for performances.
Is there any dependence between the rating of the cleanliness of the performance and the hour it takes place.
Is there any dependence between the rating of the performance and the cost of the ticket.

d. Data

The data is extracted from the system – “HHRating”. The system stores information about the cleanliness rating, customer satisfaction rating, the number of the ticket, the performance that the ticket is connected to, and the hour and the date of the performance, additionally it is known what kind of ticket they bought - the seat and the type.
In addition, data about the performance, and artists, and animals that took part in it is recorded in “Emploperf” cvs file.

2. Data sources structures

Emploperf CVS

Sheet 1 (Information about the performances in the circus, each line describes one performance, line 1 is a header row):
Column A - Performance identification number (numeric, 0 decimal precision),
Column B - Performance name (text),
Column C - Timetables of the performance – date at which performances start (in format year - month – day e.g. 2018-04-23),
Column D - Timetables of the performance – date at which performances end (in format year - month – day e.g. 2018-04-23),
Column E - Days of the performance (text, indicating the days it is currently performed - e.g. “Mondays, Thursdays”).
Note: If the timetables or days change, the data in the sheet are updated. There is always one row for one performance in the sheet.

Sheet 2 (Information about employees hired in the Circus, each line describes one employee, row 1 is a header row):
Column A - Person ID (numeric, 0 decimal precision),
Column B - Employees' PIN (PIN number),
Column C - Employee's name (text),
Column D - Employee's surname (text), in case of surname change the column is updated,
Column E - Date of birth (in format year - month – day e.g. 2018-04-23),
Column F - Education (text), in case of change the column is updated,
Column G - Date of acceptance for the position (Date in format year - month - day, e.g. 2018-04-23),
Column H - Date of end of work on the current position (Date in format year - month - day, e.g. 2018-04-23), it is not set if the employee currently works in a given position.
Column I - The main specialization of the performer (text, type of performance).

Sheet 3 (Information about animals that are owned by circus, each line describes the animal, row 1 is a header row):
Column A - The animal ID (numeric, 0 decimal precision).
Column B - The animal’s name (text).
Column C - The animal species (text).
Column D - Is available for performances (indicator if the animal can perform - yes, no).
Column E - Date the animal was acquired (Date in format year - month - day,
e.g. 2018-04-23).
Column F - Date of loss of the animal (Date in format year - month - day,
e.g. 2018-04-23).
Column G - Last veterinary visit (Date in format year - month - day, e.g. 2018-04-23).
Column H - Last cleaning of the animal (Date in format year - month - day, e.g. 2018-04-23)

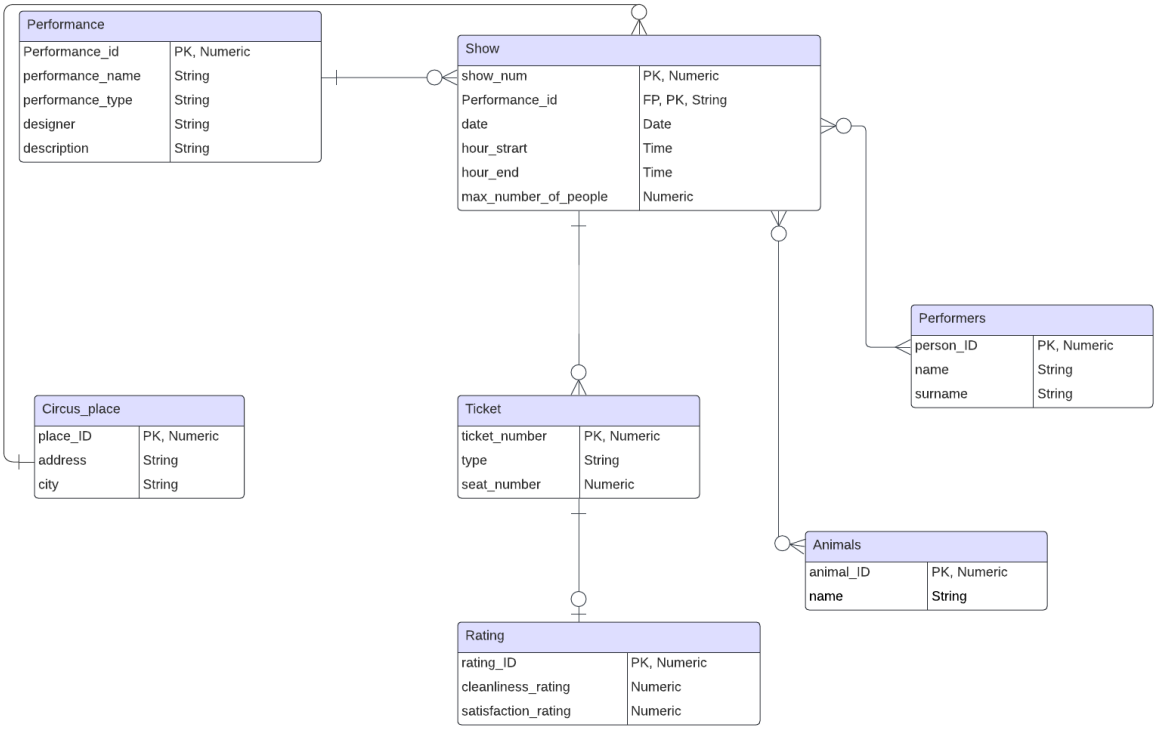
HHRating

Performance	A performance is identified by performance_id.		
	performance_id	Numeric	PK, The unique id of the performance.
	performance_name	String characters (100)	The name of the performance.

	performance_type	String characters (50)	The performance type. Allowed values: acrobatics, clown acts, animal acts, juggling, fire performances, magicians, rope acts, circus stunts, mime acts, dance performances, balancing acts, other
	designer	String characters (100)	The information about the person/people who designed the performance.
	description	String characters (1000)	The description of the performance
Show	A show is identified by show_num and performance_name.		
	show_num	Numeric	PK, The number of the showing of the specific performance.
	performance_id	Numeric	FK, part of PK
	date	Date	The date at which the show was performed.
	hour_start	Time	The time at which the show was performed.
	hour_end	Time	The time at which show ended.
	max_number_of_people	Numeric	The maximum number of people that can attend the performance.
Circus_place	Places where the circus performs, identified by place_ID		
	place_ID	Numeric	PK, The id of the place where the circus performs, added to easier distinguish.
	address	String characters (1000)	The address at which the show was performed.
	city	String characters (100)	The city at which the show was performed.
Ticket	A rating connected to the ticket bought in the circus, identified by the ticket_number.		
	ticket_number	Numeric	PK, The ticket number.

	type	String characters (20)	The type of ticket bought. Allowed values: VIP, normal, reduced.
	seat_number	Numeric	The seat number the person got.
Rating	A rating connected to the ticket bought in the circus, identified by the trating_ID.		
	rating_ID	Numeric	PK, The rating is identified by id.
	cleanliness_rating	Numeric	The cleanliness rating that customer added. Allowed values: 0-10.
	satisfaction_rating	Numeric	The satisfaction rating that the customer added. Allowed values: 0-10.
Animals	The animals performing in the performances, identified by the name.		
	animal_ID	Numeric	PK, The unique id for animals
	name	String characters (100)	The name of the animal.
Performers	The performers performing in the performances, identified by the person_ID.		
	person_ID	Numeric	PK, The performer number.
	name	String characters (100)	The name of the performer.
	surname	String characters (100)	The surname of the performer.

Relational database schema for “HHRating”



3. Scenarios of analytical problems

a. What is the effect of performers and designers on rating of performance?
   1. Compare the satisfaction rating relating to the designer of the performance.
   2. Compare cleanliness ratings depending on species of animals performing in them.
   3. Analyze if there is any correlation with the year the performers started working at and the average satisfaction rating for the performances they partook in.
   4. Compare the satisfaction rating of the performances between different performers.
   5. Analyze the average cleanliness and satisfaction rating for each performer (people and animals).
   6. Compare the performers, seeing if they excel in specific types of performances - that are not already listed as their specialization.
   7. Compare the popularity of the performers, measured through external factors (e.g., social media following), influences performance ratings.
   8. Compare the relationship between employee performance ratings and their reported satisfaction with work-life balance.

b. What influences the number of people doing the surveys?
   1. Analyze the relationship between the type of ticket (VIP, normal, reduced) and survey participation.
   2. Analyze whether types of performances with great reputations (higher ratio of the satisfied ratings) have higher participation rates compared to the performances with worse reputation.

3. Compare whether the time of day the performance takes place affects the survey participation rate.
4. Identify whether the uniqueness of certain events motivates visitors to provide feedback.
5. Analyze survey participation based on where the audience is sitting.
6. Compare the number of people participating in the surveys depending on the species of the animal performing.
7. Analyze whether good circus maintenance (cleanliness, greater audience satisfaction) influences participation in surveys.
8. Analyze whether the frequency of veterinary visits for performing animals influences audience participants in the survey. Assess whether well-cared-for animals lead to increased audience engagement.

4. Data needed for analytical problems

a. Analytical problem: What is the effect of performers on rating of performance?

1. Compare the satisfaction rating relating to the designer of the performance.
   - designer of the performance - *HHRating*, table *Performance*, column *designer,*
   - average satisfaction score - calculated average satisfaction rating for the performence from *HHRating*, table *Rating*, column *satisfaction_rating*.
2. Compare cleanliness ratings depending on species of animals performing in them.
   - animals performing - *HHRating*, table *Animals*, column *animal_ID* and *Emploperf cvs* sheet 3, Column A - the animals ID,
   - animal species - *Emploperf cvs*, sheet 3, Column C - the animal species,
   - average cleanliness rating - calculated from *HHRating*, table *Rating*, column *cleanliness_rating.*
3. Analyze if there is any correlation with the year the performers started working at and the average satisfaction rating for the performances they partook in.
   - performers performing - *HHRating*, table *Performers*, column *person_ID,*
   - performances that they partook in - *HHRating*, table *Performance*, column *performance_id*
   - average satisfaction score - calculated average satisfaction rating for the performance from *HHRating*, table *Rating*, column *satisfaction_rating*,
   - date they started working at - *Emploperf cvs*, sheet 2, Column G - date of acceptance for the position.
4. Compare the satisfaction rating of the performances between different performers.
   - performers performing - *HHRating*, table *Performers*, column *person_ID*,
   - performances that they partook in - *HHRating*, table *Performance*, column *performance_id*,
   - average satisfaction score - calculated average satisfaction rating for the performance from *HHRating*, table *Rating*, column *satisfaction_rating*.
5. Analyze the average cleanliness and satisfaction rating for each performer (people and animals).
   - animal performing - *HHRating*, table *Animals*, column *animals_ID*,
   - person performing - *HHRating*, table *Performers*, column *person_ID*,
   - average cleanliness rating - calculated average cleanliness rating for the performance from *HHRating*, table *Rating*, column *cleanliness_rating*,
   - average satisfaction rating - calculated average satisfaction rating for the performance from *HHRating*, table *Rating*, column *satisfaction_rating*.

6. Compare the performers, seeing if they excel in specific types of performances - that are not already listed as their specialization.

- performer - *HHRating*, table *Performers*, column *person_ID*,
- performance type - *HHRating*, table *Performance*, column *performance_type*,
- performers specialization - *Emploperf cvs*, sheet 2, Column I - The main specialization of the performer,
- average satisfaction rating - calculated average satisfaction rating for the performance from *HHRating*, table *Rating*, column *satisfaction_rating*.

7. Compare the popularity of the performers, measured through external factors (e.g., social media following), influences performance ratings.

- performer - *HHRating*, table *Performers*, column *person_ID*,
- average satisfaction rating - calculated average satisfaction rating for the performance from *HHRating*, table *Rating*, column *satisfaction_rating*,
- popularity of the performer - not visible in our system, measured through external factors, the proposal of acquiring such an information:
    - social media following,
    - the number of views of performances on YouTube.

8. Compare the relationship between employee performance ratings and their reported satisfaction with work-life balance.

- performer - *HHRating*, table *Performers*, column *person_ID*,
- average satisfaction rating - calculated average satisfaction rating for the performance from *HHRating*, table *Rating*, column *satisfaction_rating.*

It is not possible to build a BI system to solve this analytical question without expanding the rating business activity. It is suggested to introduce the employee satisfaction rating, in which they would each month rate their work with the minimum of the following criterions:

- On the scale from 0-10 how would you rate your work-life balance?
- On the scale from 0-10 how would you rate your satisfaction with this month's performances?

Survey results are automatically uploaded to the survey system and entered into the cvs file.
Column A - Performers ID (numeric)
Column B - Date of the survey.
Column C - Rating of the work-life balance (0-10).
Column D - Rating of the satisfaction (0-10).

b. Analytical problem: What influences the number of people doing the surveys?

1. Analyze the relationship between the type of ticket (VIP, normal, reduced) and survey participation.

- ticket type - *HHRating*, table *Ticket*, column *type*,
- the percent of people with certain ticket types participating in surveys - calculated dividing them by different ticket types and then dividing the number of ratings from *HHRating*, table *Rating*, column *rating_ID* by the number of tickets from *HHRating*, table *Ticket*, column *ticket_number*.

2. Analyze whether types of performances with great reputations (higher ratio of the satisfied ratings) have higher participation rates compared to the performances with worse reputation.

- performance type - *HHRating*, table *Performances*, column *performance_type*,
- the good reputation of the performance - calculated from dividing the number of 10-8 satisfaction ratings from *HHRating*, table *Rating*, column *satisfaction_rating*

by the number of people that went to the performance from *HHRating*, table *Rating*, column *rating_ID*,
- the percent of people participating in the survey depending on type of performance - calculated from the number of ratings connected to the type of performance - *HHRating*, table *Rating*, column *rating_ID* divided by the number of tickets bought connected to the type of performance - *HHRating*, table *Ticket*, column *ticket_number*.

3. Compare whether the time of day the performance takes place affects the survey participation rate.
    - hour of the performance start - *HHRating*, table *Show*, column *hour_start*,
    - performance type - *HHRating*, table *Performance*, column *performance_type*,
    - the percent of people participating in the surveys - calculated from the number of ratings connected to the hour of performance - *HHRating*, table *Rating*, column *rating_ID* divided by the number of tickets bought connected to hour of performance - *HHRating*, table *Ticket*, column *ticket_number*.
4. Identify whether the uniqueness of certain performances motivates visitors to provide feedback.
    - the number of the performances of the certain type - calculated from *HHRating*, table *Performance*, column *performance_type*,
    - the percent of people participating in the surveys - calculated from the number of ratings - *HHRating*, table *Rating*, column *rating_ID* divided by the number of tickets bought - *HHRating*, table *Ticket*, column *ticket_number* (connected to the performance type).
5. Analyze survey participation based on where the audience is sitting.
    - seat number - HHRating, table Ticket, column seat_number,
    - the percent of people participating in the surveys - calculated from the number of ratings - *HHRating*, table *Rating*, column *rating_ID* divided by the number of tickets bought - *HHRating*, table *Ticket*, column *ticket_number* (connected to seat number).
6. Compare the number of people participating in the surveys depending on the species of the animal performing.
    - animals performing - *HHRating*, table *Animals*, column *animal_ID* and *Emploperf cvs*, sheet 3, Column A - the animal ID,
    - animal species - *Emploperf cvs*, sheet 3, Column C - the animal species,
    - the percent of people participating in the surveys - calculated from the number of ratings - *HHRating*, table *Rating*, column *rating_ID* divided by the number of tickets bought - *HHRating*, table *Ticket*, column *ticket_number* (connected to the performances by the animals that participate).
7. Analyze whether good circus maintenance (cleanliness, greater audience satisfaction) influences participation in surveys.
    - average cleanliness rating - calculated average cleanliness rating for the performance from *HHRating*, table *Rating*, column *cleanliness_rating*,
    - average satisfaction rating - calculated average satisfaction rating for the performance from *HHRating*, table *Rating*, column *satisfaction_rating*,
    - the percent of people participating in the surveys - calculated from the number of ratings - *HHRating*, table *Rating*, column *rating_ID* divided by the number of tickets bought - *HHRating*, table *Ticket*, column *ticket_number.*