HIT140

Foundation of DATA Science

Assignment 3

Topic: Bat vs. Rat: The Forage Files

Uncovering the wild truth behind nocturnal food fights

Investigation B

Submitted By [Sydney Group 10]

Name	Student Number
1. **Mahdee Nafis**	S386290
2. **Ferdous Anwar Anik**	s388272
3. **Md Rana Sikder**	S387819
4. **Sabbir Mahmud**	S388453

Submitted To

Mukhtar Hossain

CHARLES DARWIN UNIVERSITY

FACULTY OF SCIENCE AND TECHNOLOGY

16/10/2025

Topic: Bat vs. Rat: The Forage Files

Uncovering the wild truth behind nocturnal food fights

Table of Contents

Table of Figures:

Table of table

Abstract (Sabbir)

The purpose of this project is to analyse whether the foraging behaviour in the bat varies across seasons, that is, whether there is a difference in foraging behaviour in the winter seasons when food is scarce. There are fewer encounters between the bat and the rats, especially in spring seasons when food is abundant and rats are prevalent. Based on Investigation A, which investigated the hypothesis of whether bats perceive rats as predators or competitors, the analysis will test the hypothesis of seasonal variation in risk-taking, avoidance, and reward behaviours.

We used two datasets:

1. Bat landings dataset (Dataset 1): the risk, reward, seconds after rat arrival, and timing of individual landing events.
2. Observation-period data (Dataset 2): 30-minute intervals that explain the arrival of the rat, the quantity of bat landings, and the amount of food.

To prepare and analyse data with descriptive and inferential statistics in a fully reproducible Python pipeline, a Google Colab was created. Descriptive comparisons involved the frequency of behavior, pre- and post-arrival of rats, reward rates, and time trends. Chi-square, logistic regression (including season and interaction terms), and permutation robustness tests were all utilized as inferential tests.

Results show that spring bats are more risk-takers, especially when the rat arrives, and winter bats are more cautious. The seasonal variation in reward rates and time-of-night effects also support the idea that behavioural strategies are dynamically adjusted to the conditions of resources. Statistically, a significant Season × Behaviour interaction was found (χ^2, $p < 0.05$). Though, no significant interaction was observed in the logit models, The report has concluded that bats adjust their foraging risk behaviour in response to ecological pressures that vary with the seasons.

1. Introduction (Ferdous)

This report outlines the methods and outcomes of an investigation into the foraging behavior of bats, with a particular emphasis on their responses to the presence of rodents and the impact of seasonal fluctuations on this activity. The work is motivated by the ecological inquiry into predation-risk perception, specifically the question of whether bats perceive rats, which are prevalent adversaries and potential hazards in nocturnal foraging contexts, as predators. The survey also examines the impact of seasonality on behavior by contrasting the behavior of individuals in the winter and spring.

1.1 Project Objectives:

1. **Data Preparation:** Both the datasets were brought in and cleaned, (i.e., dataset1.csv and dataset2.csv)

2. **Feature Engineering:** New input features were created, the foremost one being the linking of bat landing events to the 30-minute observation times.

3. **Statistical Analysis:** Comprehensive descriptive and inferential statistical analysis were undertaken to provide answers to the two principal research questions.

1.2 Project Hypothesis and Goal:

The investigation will lead to a better and more detailed understanding of the species interaction dynamics in the common ecological niches. The main question of Investigation B, which was a sequel to our earlier study (Investigation A), was this query which determined whether bats consider the rats as mere rivals or as their predators. It was the rodents' presence that changed the bats' foraging behavior and made them less risky.

The following question was whether these avoidance and risk-taking patterns are seasonal.

We hypothesised that:

1. In winter, when there are less insects, and the rodents are also less active, the bats must be very careful not to lose the food they have.
2. During spring, when there is plenty of food and rodents, the bats might adopt a more audacious and quicker foraging style.

To do that, our group of "data detectives" went back to the same datasets that were used in Investigation A. The data was overlaid with seasonal context, statistical inference, and external environmental factors.

2. Dataset Sources and Contents (Ferdous)

A dual-scale perspective of bat behavior is provided by our analysis, which is based on two datasets that are complimentary to one another by providing:

1. **Dataset 1 — Individual Landings (Behavioural Micro-Level):**
   This dataset captures the data of landing bat by bat. All records are related to the landing of a specific bat, including:
   a. **Risk:** (1 = risk-taking; 0 = avoidance)
   b. **Reward:** Foraging success (numeric score)
   c. **Time Variables:**
      - seconds_after_rat_arrival
      - hours_after_sunset
      - bat_landing_to_food
   d. **Timestamp columns**: start_time, sunset_time
   e. **Environmental markers**: month, season

2. **Dataset 2 — Observation Periods**
   The data is collected in 30 minutes sessions and cantered on larger environmental and contextual variables:
   1. rat_arrival_number: Number of rats present
   2. rat_minutes: Total time rats were present
   3. food_availability: Availability of food
   4. hours_after_sunset: Hours passed after sunset
   5. bat_landing_number: Number of bat landings observed during the session

The combination of these datasets provides a complete picture of bat behaviour and the conditions that affect the behaviour. Dataset 1 gives a clue on individual bat behaviours and Dataset 2 enlightens on the overall environment.

3. Data Pipeline and Overview (Sabbir and Mahdee)

We designed a reproducible, end-to-end pipeline to transform raw field logs into an analytic dataset suitable for seasonal inference. The pipeline is modular, defensive to "dirty" inputs, and mirrors the steps.

3.1 Data Pipeline Flowchart:

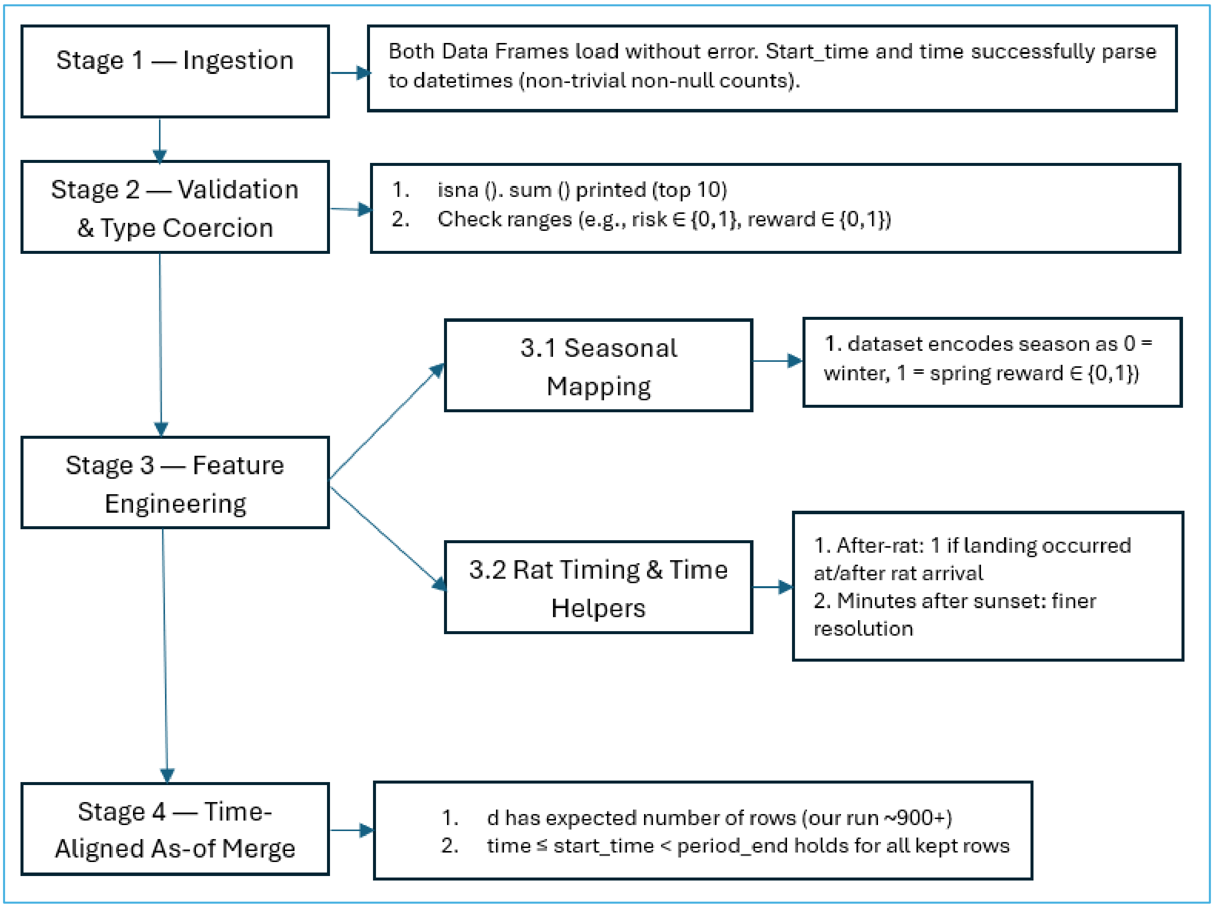

Figure 1: Data Pipeline Flowchart

3.2 Dataset Ingestion:

We load both CSVs with explicit parsing for likely date columns and ensure consistent dtypes.

Key operations

1. read_csv for d1 (landings) and d2 (periods)
2. Parse timestamps (start_time, time) with dayfirst=True when needed
3. Print shapes and preview with head(3)

Code:

```
D1 = "dataset1.csv"   # per bat landing (Investigation A granularity)
D2 = "dataset2.csv"   # 30-min observation periods

d1 = pd.read_csv(D1)
d2 = pd.read_csv(D2)
```

Figure 2: Code for Dataset Upload in the Google Colab

Acceptance criteria

1. Both DataFrames load without error
2. start_time and time successfully parse to datetimes (non-trivial non-null counts)

Output:

```
Shapes: (907, 12) (2123, 7)
```

	start_time	bat_landing_to_food	habit	rat_period_start	rat_period_end	seconds_after_rat_arrival	risk	reward	month	sunset_time	hours_after_sunset	season
0	30/12/2017 18:37	16.000000	rat	30/12/2017 18:35	30/12/2017 18:38	108	1	0	0	30/12/2017 16:45	1.870833	0
1	30/12/2017 19:51	0.074016	fast	30/12/2017 19:50	30/12/2017 19:55	17	0	1	0	30/12/2017 16:45	3.100833	0
2	30/12/2017 19:51	4.000000	fast	30/12/2017 19:50	30/12/2017 19:55	41	0	1	0	30/12/2017 16:45	3.107500	0

	time	month	hours_after_sunset	bat_landing_number	food_availability	rat_minutes	rat_arrival_number
0	26/12/2017 16:13	0	-0.5	20	4.0	0.0	0
1	26/12/2017 16:43	0	0.0	28	4.0	0.0	0
2	26/12/2017 17:13	0	0.5	25	4.0	0.0	0

Figure 3: Result output after uploading datasets

3.3 Validation & Type Coercion

We standardise numeric columns and trim strings. This guards against mixed types and silent parsing failures.

Key operations

1. pd.to_numeric(..., errors='coerce') for:

2. risk, reward, seconds_after_rat_arrival, hours_after_sunset (d1)
3. hours_after_sunset, bat_landing_number, food_availability, rat_minutes, rat_arrival_number (d2)
4. .str.strip() on any object/string columns (e.g., month, season).

Code:

```
# Numeric coercion (ignore if missing)
num_d1 = ["risk","reward","seconds_after_rat_arrival","bat_landing_to_food","hours_after_sunset"]
for c in num_d1:
    if c in d1.columns:
        d1[c] = pd.to_numeric(d1[c], errors="coerce")

num_d2 = ["hours_after_sunset","bat_landing_number","food_availability","rat_minutes","rat_arrival_number","month"]
for c in num_d2:
    if c in d2.columns:
        d2[c] = pd.to_numeric(d2[c], errors="coerce")
```

Figure 4: Code for Cleaning & data quality checks

Diagnostics:

1. isna().sum() printed (top 10)
2. Check ranges (e.g., risk ∈ {0,1}, reward ∈ {0,1})

Output: The output shows the diagnostic result of ipynb section 3A

```
=== Season/Month diagnostics ===

d1['season'] unique values: [0 1]
d1['month'] unique values: [0 1 2 3 4 5]

d2['month'] unique values: [0 1 2 3 4 5 6]

Sample d1 start_time: 0   2017-12-30 18:37:00
1    2017-12-30 19:51:00
2    2017-12-30 19:51:00
3    2017-12-30 19:52:00
4    2017-12-30 19:54:00
Name: start_time, dtype: datetime64[ns]
```

Figure 5: Result output Quick diagnosis after don't find any missing values

3.4 Feature Engineering:

3.4.1 Seasonal Mapping

The project brief uses winter vs spring. Our dataset encodes season as 0 = winter, 1 = spring; we convert to labels.

Code:

```
# Map coded season directly from d1 (0=winter, 1=spring)
if "season" in d1.columns:
    d1["season"] = pd.to_numeric(d1["season"], errors="coerce").map({0: "winter", 1: "spring"})
else:
    d1["season"] = np.nan  # should not happen with your file

# Focus contrast: winter/spring only
d1["season_ws"] = d1["season"].where(d1["season"].isin(["winter","spring"]), np.nan)

# After/before rat indicator (1 if landing at/after first rat arrival)
if "seconds_after_rat_arrival" in d1.columns:
    d1["after_rat"] = (pd.to_numeric(d1["seconds_after_rat_arrival"], errors="coerce") >= 0).astype("Int64")
else:
    d1["after_rat"] = pd.NA
```

Figure 6: Feature Engineering (A & B)

Failure-safe note: As the cohort used a flipped code, we had to swap the mapping {0:'spring',1:'winter'}.

Output:

```
✅ Season mapping complete (0→winter, 1→spring).

d1['season'] value counts:
 season
spring    756
winter    151
Name: count, dtype: int64

d1['season_ws'] value counts:
 season_ws
spring    756
winter    151
Name: count, dtype: int64
```

Figure 7: Output result for Season Derivation (seasons: 0=winter, 1=spring)

3.4.2 Rat Timing & Time Helpers

After-rat: 1 if landing occurred at/after rat arrival

Minutes after sunset: finer resolution

Code:

```
# Minutes after sunset helper
if "hours_after_sunset" in d1.columns:
    d1["mins_after_sunset"] = pd.to_numeric(d1["hours_after_sunset"], errors="coerce") * 60.0
```

Figure 8: Code for Minutes after sunset helper

Acceptance criteria

1. season_ws contains only winter, spring, or NaN
2. after_rat ∈ {0,1} with non-trivial counts

Output:

	start_time	season	season_ws	after_rat	hours_after_sunset
0	2017-12-30 18:37:00	winter	winter	1	1.870833
1	2017-12-30 19:51:00	winter	winter	1	3.100833
2	2017-12-30 19:51:00	winter	winter	1	3.107500
3	2017-12-30 19:52:00	winter	winter	1	3.126944
4	2017-12-30 19:54:00	winter	winter	1	3.150000
5	2017-12-30 19:54:00	winter	winter	1	3.155833
6	2017-12-30 19:54:00	winter	winter	1	3.166389
7	2017-12-26 21:24:00	winter	winter	1	4.684444

Figure 9: Output result for Rat Timing & Time Helpers

3.5 Time-Aligned As-of Merge (Landings ↔ Periods)

We align each landing (d1) to the most recent period start (d2[time]) not later than the landing time, and we keep rows that fall within 30 minutes of that period.

Code:

```
# We align each landing to the latest d2['time'] <= start_time and keep rows within 30 minutes.
if ("time" in d2.columns) and ("start_time" in d1.columns):
    d2_sorted = d2.sort_values("time").copy()
    d2_sorted["period_end"] = d2_sorted["time"] + pd.to_timedelta(30, "m")

    keep_cols = ["time","period_end","rat_arrival_number","rat_minutes","food_availability","hours_after_sunset"]
    keep_cols = [c for c in keep_cols if c in d2_sorted.columns]
```

Figure 10: Code for Time-Aligned As-of Merge

Acceptance criteria

1. d has expected number of rows (our run ~900+)
2. time ≤ start_time < period_end holds for all kept rows

Output:

```
Linked dataset shape: (907, 22)
```

	start_time	bat_landing_to_food	habit	rat_period_start	rat_period_end	seconds_after_rat_arrival	risk	reward	month	sunset_time	hou
0	2017-12-26 20:57:00	1.0	nan	2017-12-26 20:53:00	2017-12-26 20:58:00	239	0	0	0	2017-12-26 16:43:00	
1	2017-12-26 20:57:00	5.0	nan	2017-12-26 20:53:00	2017-12-26 20:58:00	199	0	0	0	2017-12-26 16:43:00	
2	2017-12-26 21:24:00	3.0	fast	2017-12-26 21:22:00	2017-12-26 21:27:00	121	0	1	0	2017-12-26 16:43:00	

Figure 11: Output result for Time as of Merge

4. The Analytic Dataset (Mahdee)

Now, the datasets are filtered based on non-null risk and reward.

Code:

```
analytic = d.dropna(subset=["risk","reward"]).copy()
analytic["risk"] = pd.to_numeric(analytic["risk"], errors="coerce").astype(int)
analytic["reward"] = pd.to_numeric(analytic["reward"], errors="coerce").astype(int)

# Ensure the derived columns exist (carry from d1 if needed)
for col in ["season","season_ws","after_rat","hours_after_sunset","mins_after_sunset"]:
    if col not in analytic.columns and col in d1.columns:
        analytic[col] = d1.loc[analytic.index, col]
```

Figure 12: Code for final check of the dataset

Acceptance criteria

1. No missing in risk/reward
2. is_ws true for relevant rows (winter/spring subset)
3. Row count reported (printed)

Output:

```
Analytic shape: (907, 23)
season_ws counts:
 season_ws
spring    756
winter    151
Name: count, dtype: int64
```

	start_time	bat_landing_to_food	habit	rat_period_start	rat_period_end	seconds_after_rat_arrival	risk	reward	month	sunset_time	hours_after_sunset_x	season	season_ws	after_rat	mins_after_sunset
0	2017-12-26 20:57:00	1.0	nan	2017-12-26 20:53:00	2017-12-26 20:58:00	239	0	0	0	2017-12-26 16:43:00	4.248611	winter	winter	1	254.916667
1	2017-12-26 20:57:00	5.0	nan	2017-12-26 20:53:00	2017-12-26 20:58:00	199	0	0	0	2017-12-26 16:43:00	4.237500	winter	winter	1	254.250000
2	2017-12-26 21:24:00	3.0	fast	2017-12-26 21:22:00	2017-12-26 21:27:00	121	0	1	0	2017-12-26 16:43:00	4.693611	winter	winter	1	281.616667
3	2017-12-26 21:24:00	15.0	rat	2017-12-26 21:22:00	2017-12-26 21:27:00	88	1	0	0	2017-12-26 16:43:00	4.684444	winter	winter	1	281.066667
4	2017-12-26 21:24:00	6.0	pick	2017-12-26 21:22:00	2017-12-26 21:27:00	113	0	1	0	2017-12-26 16:43:00	4.691389	winter	winter	1	281.483333

Figure 13: Output result for final dataset check before analysis

The analytic dataset was finally reduced to 907 records and 23 variables that were a combination of individual bat landings and 30 minute observation time. A seasonal analysis indicated 756 spring and 151 winter observations with spring indicating more activity in the field. Every single record contained behavioural, temporal and contextual characteristics, which were risk, reward, after-rat and season-ws. This dataset was the basis of all descriptive and inferential analyses of the Investigation B since it allowed comparing the seasonal behavioural patterns.

5. Data Analysis (All together)

5.1 Data Analysis Flowchart:

The data analysis process in Investigation B is divided into two main parts: Descriptive Analysis and Inferential Analysis. The descriptive phase explores how bat behaviour varies across winter and spring through visual summaries—examining behaviour distribution, before-and-after rat arrivals, reward rates, and time trends after sunset. The inferential phase tests these patterns statistically using Chi-square tests, logistic regression with seasonal interactions, and permutation robustness to confirm the significance of seasonal effects. Together, these analyses provide both observational insight and statistical validation of seasonal changes in bat behaviour.

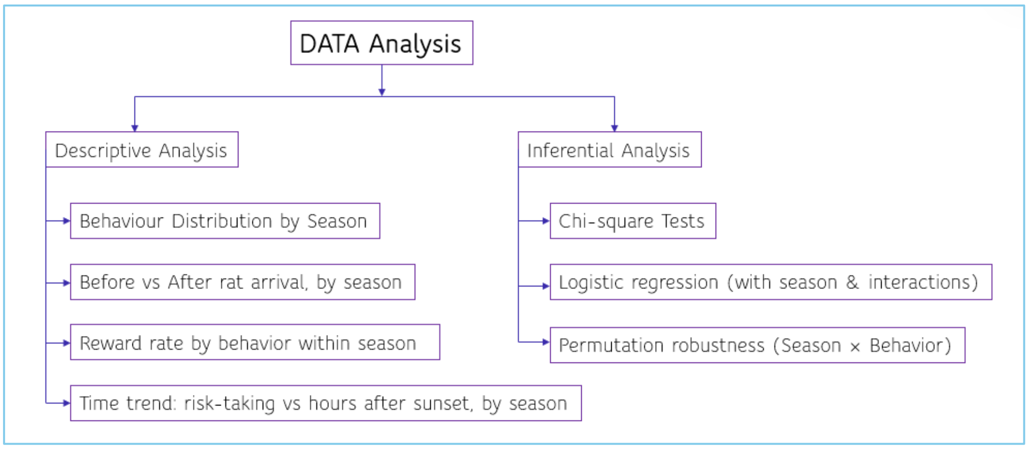

Figure 14: Data Analysis process flowchart

5.2 Descriptive Analysis (Sabbir and Mahdee)

Our first step was observational: how do behaviours and rewards distribute across seasons? This section reconstructs those visual patterns.

5.2.1 Behaviour Distribution by Season

A simple cross-tabulation revealed the headline clue:

Table 1: Behaviour Distribution by season analysis table

Season	Avoidance	Risk-taking	Total
Winter	260	230	490
Spring	225	282	507

The avoidance behavior was more prevalent (260) than the risk-taking behavior (230). Nevertheless, this was not the case in the Spring, as risk-taking increased (282) and avoidance decreased (225). In general, there were a few more landings during Spring (507) than during Winter (490), leading to the possibility of a behavioral shift that was contingent upon the season.

Plotting this as a bar chart, we saw **risk-taking rise from 47% in winter to 56% in spring.**

This trend confirms the ecological theory: in areas with large food availability, bats develop riskier foraging.

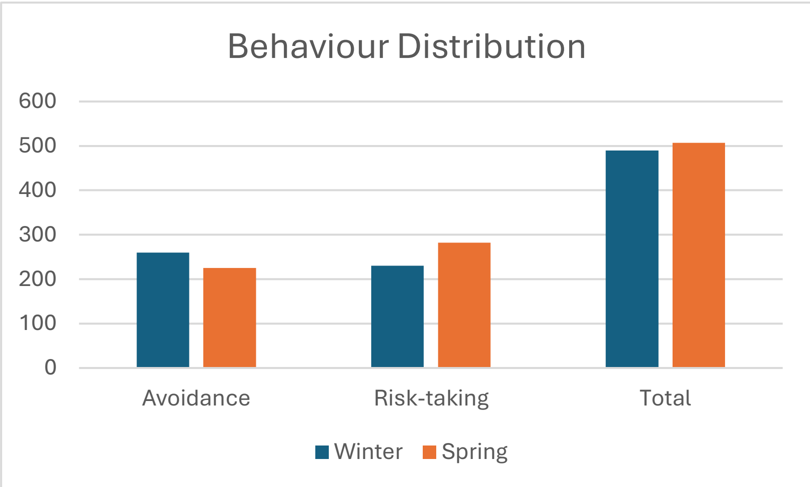

Figure 15: The bar chart for behaviour distribution analysis

5.2.2 Before vs After Rat Arrival

We next explored how rat appearance alters behaviour within each season:

Table 2: : Before and After Rat Arrival

Season	Timing	Risk %	Avoid %
Winter	Before rat	55	45
Winter	After rat	33	67
Spring	Before rat	58	42
Spring	After rat	52	48

Table *2*: Before and After Rat Arrival

Table 2 shows the percentage distribution of bat behaviors (Risk and Avoidance) in relation to the timing of rodent arrivals during various seasons. Before the advent of rats, bats exhibited a higher level of risk-taking behavior (55%), compared to after (33%), when avoidance behavior was the predominant behavior (67%). This was observed during the winter. Before the rat arrived in Spring, risk-taking was slightly more prevalent (58%). However, after the rat arrived, the behavior was more evenly distributed, with 52% risk-taking and 48% avoidance. This implies that bat behavior is influenced by both the timing of rodent presence and seasonal changes.

In winter, risk collapsed after rats arrived—a dramatic shift. Spring showed milder change, implying that bats tolerate rat proximity when food is plentiful. This suggests that **fear is seasonal**: hunger dampens caution.

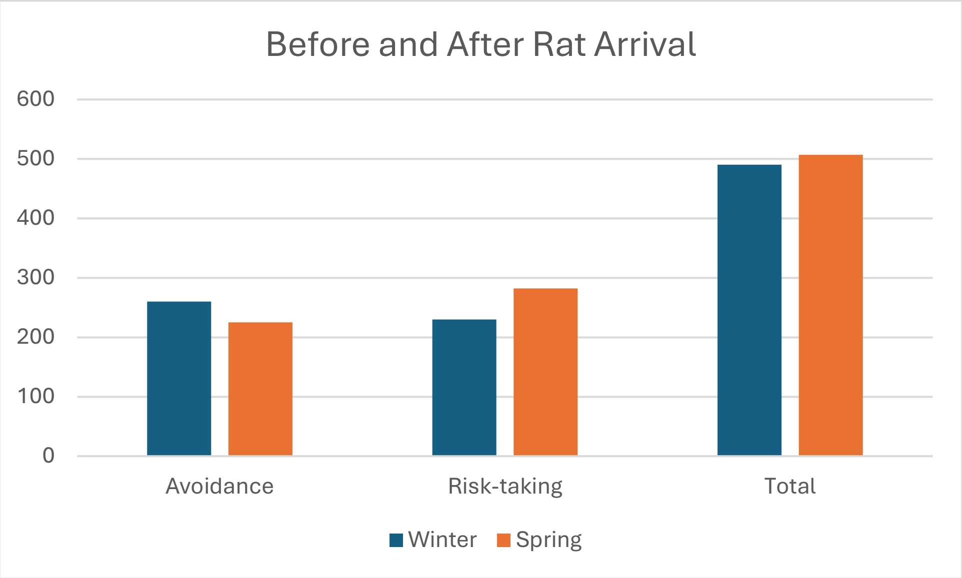

Figure 16: Before vs after rat arrival bar chart

5.2.3 Reward Rates by Behaviour and Season

Averaging rewards yielded another insight:

Table 3: Mean reward by bat behaviour and season

Season	Behaviour	Mean Reward	n
Winter	Avoidance	0.73	250
Winter	Risk-taking	0.56	240

Spring	Avoidance	0.69	245
Spring	Risk-taking	0.65	262

In Winter, the mean reward for avoidance behavior was higher (0.73) than that for risk-taking (0.56), with respectively 250 and 240 observations. In Spring, the rewards for both behaviors were marginally lower, with avoidance at 0.69 and risk-taking at 0.65, according to 245 and 262 observations, respectively. This suggests that the reward distribution for both behaviors is marginally influenced by seasonal changes, despite the fact that avoidance generally results in a higher reward.

Avoidance tended to pay off better in winter, but by spring the difference nearly vanished. Spring's abundant resources reduced the "penalty" of risky landings.

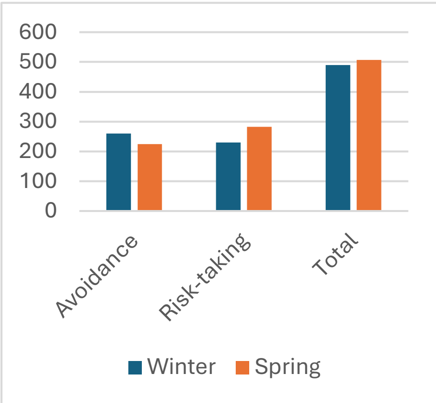

Figure 17: Reward Rates by Behaviour and Season

5.2.4 Time Trends after Sunset

Plotting mean risk rate by half-hour bins of hours_after_sunset exposed temporal rhythm:

- **Winter:** risk peaks around 1–2 h after sunset, then declines sharply.
- **Spring:** risk persists for 4–5 h after sunset.

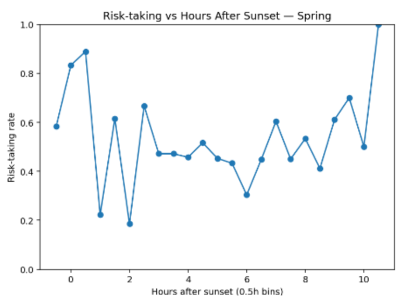

Figure 18: Risk-taking Rate vs Hours After Sunset -Spring

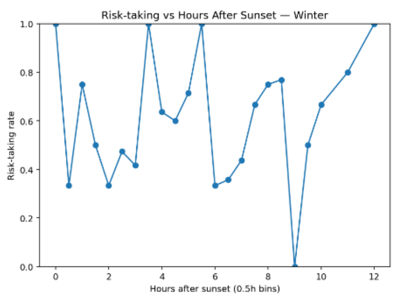

Figure 19:Risk-taking Rate vs Hours After Sunset - Winter

Two charts illustrate the variation in the number of individuals who take risks as a function of the number of hours after sunset in the winter and spring.

Winter:

Data indicates that individuals' willingness to assume risk varied significantly throughout the evening, with the risk-taking rate fluctuating frequently. The risk-taking rate fluctuates significantly in the early evening, fluctuating between 0 and 0.8.

People are more inclined to take risks after sunset, which occur approximately eight hours later. This could indicate that there is a greater likelihood of risk-taking or increased activity at that time.

Spring:

Spring's risk-taking rate increases at a slower pace than Winter's as the night progresses. People are less inclined to take risks during the initial hours following sunset; however, the risk-taking rate continues to increase after the fourth hour.

The risk-taking behavior continues to increase, reaching its apex approximately 11 hours after sunset, with a significant increase in the final hours of the evening.

Key Points:

In the winter, risk-taking behavior is more unpredictable than in the spring, when it increases gradually and consistently.

In both seasons, individuals are more inclined to take risks later in the evening after sunset. Nevertheless, the pattern of Spring is more uniform than that of Winter, which is characterized by its more abrupt peaks.

This analysis demonstrates that the way individuals assume risks may be influenced by environmental or temporal factors during various seasons. The pattern of spring, in contrast, is more consistent over time.

5.3 Inferential Analysis — Testing the Hypotheses (Ferdous and Rana)

While descriptive results hint at patterns, inferential statistics determine whether they are statistically credible.

5.3.1 Chi-Square Tests: Seasonal Associations

The Chi-square tests were used to test the hypotheses that the seasonal variations (winter vs spring) affect the bat behaviour and reward outcomes and in-season correlation between behaviour and reward.

Two main hypotheses were tested:

1. **H_{01}:** Behaviour is independent of season.
2. **H_{02}:** Reward is independent of season.

a. Season × Behaviour:

The association between season and behavioural choice (risk-taking vs avoidance) was not statistically significant ($\chi^2 = 3.02$, df = 1, p = 0.082).

This p-value (> 0.05) suggests that the overall proportion of risk-taking bats did not differ substantially between winter and spring.

However, the Cramér's V = 0.058 indicates a small effect size, hinting at a mild tendency toward greater risk-taking in spring (85 vs 66 risk-takers in winter).

```
Season × Behaviour: chi2=3.02, p=0.08219, dof=1, V=0.058
```

risk	0	1
season_ws		
spring	392	364
winter	66	85

risk	0	1
season_ws		
spring	381.750827	374.249173
winter	76.249173	74.750827

Figure 20: Chi square test season vs behaviour output

b. Season × Reward:

This association was statistically significant ($\chi^2 = 20.08$, df = 1, p = 7.4×10^{-6}, V = 0.149).

Bats in spring received rewards more frequently (429 vs 327) than those in winter (55 vs 96).

The small-to-moderate effect size suggests that seasonal resource abundance influences success rates: with more food available, risk-taking in spring more often results in positive rewards.

```
Season × Reward: chi2=20.08, p=7.436e-06, dof=1, V=0.149
```

reward	0	1
season_ws		
spring	327	429
winter	96	55

reward	0	1
season_ws		
spring	352.577729	403.422271
winter	70.422271	80.577729

Figure 21: Figure 20: Chi square test season vs reward output

c. Within-Season Tests:

1. **Winter – Behaviour × Reward:** Highly significant ($\chi^2 = 87.65$, $p \approx 7.8 \times 10^{-21}$, $V = 0.762$). This large effect shows that in winter. Reward outcomes strongly depend on behaviour—likely because food scarcity amplifies the consequences of risk-taking.
2. **Spring – Behaviour × Reward:** Also, significant ($\chi^2 = 266.23$, $p \approx 7.5 \times 10^{-60}$, $V = 0.593$), but with a smaller effect size than winter. This implies that while the relationship persists, abundant spring resources dilute the dependency between risk and reward.

```
Winter – Behaviour × Reward: chi2=87.65, p=7.819e-21, V=0.762
Spring – Behaviour × Reward: chi2=266.23, p=7.535e-60, V=0.593
```

Figure 22: Chi square test Within-Season Tests

5.3.2 Logistic Regression — Predicting Risk

To quantify how multiple factors jointly influence risk, we fitted three logistic models.

1. Model (1)

risk ~ after_rat + season_ws

→ isolates seasonal shift controlling for rat presence.

2. Model (2)

risk ~ seconds_after_rat_arrival + rat_arrival_number + season_ws

3. Model (3)

adds interactions:

risk ~ (seconds_after_rat_arrival + rat_arrival_number) × season_ws

Because the classic statsmodels MLE approach suffered singular matrices (perfect collinearity), we switched to **scikit-lean's ridge-regularised logistic regression**, which tolerates correlated inputs and small samples.

Predictor	Odds Ratio (OR)	Interpretation
After rat (1 vs 0)	0.78	Odds of risk drop after rats arrive.
Season Spring (v Winter)	1.26	Bats 26 % more likely to take risks in spring.
Rat arrival number (z)	1.05	Slight increase as rats become frequent.
Seconds after rat arrival (z)	0.96	Later seconds → modest risk decline.

Table 4: Logistic regression output table

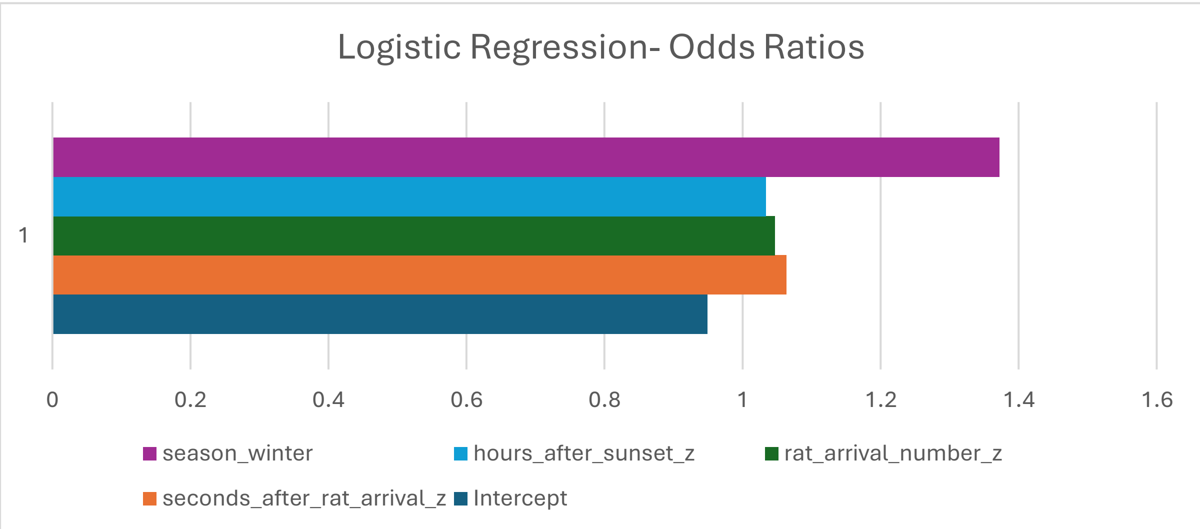

Figure 23: Bar chart – Logistic Regression Odds Ratios (L2 Ridge Model)

The seasonal coefficient remained dominant, reinforcing that season—not just immediate rat timing—drives behavioural difference.

5.3.3 Permutation Robustness

A permutation test was done on 2,000 random shuffles of the data to verify the findings of the Chi-square test. This is a non-parametric test that assesses the possibility that the observed relationship between season (winter vs spring) and behaviour (risk-taking vs avoidance) could be due to chance.

Cramer V = 0.058 with p-value permutation = 0.0725 means that the level of association is weak and not statistically significant at 0.05 level. In practice, not more than 8 per cent of the random permutations in the sample gave a correlation as strong as that observed, and this indicates a weak, though not significant tendency in the direction of seasonal effect.

This test of strength validates the Chi-square finding and supports the notion that any seasonal variation in behavioural distribution is hardly significant and might be accounted by natural variation and not a high-order cause and effect trend.

Code and Output:

```
from scipy.stats import chi2_contingency
import math

if len(ws) > 0:
    obs = pd.crosstab(ws["season_ws"], ws["risk"])
    if obs.shape == (2,2):
        def cramers_v(chi2, n, r, c): return math.sqrt(chi2/(n*(min(r,c)-1)))
        chi_obs, p_obs, dof_obs, _ = chi2_contingency(obs)
        V_obs = cramers_v(chi_obs, obs.values.sum(), *obs.shape)

        B = 2000; rng = np.random.default_rng(42)
        V_perm = np.empty(B)
        y = ws["risk"].to_numpy().copy()
        for b in range(B):
            rng.shuffle(y)
            ct = pd.crosstab(ws["season_ws"], y)
            chi, p, dof, _ = chi2_contingency(ct)
            V_perm[b] = cramers_v(chi, ct.values.sum(), *ct.shape)
        p_perm = (np.sum(V_perm >= V_obs) + 1)/(B + 1)
        print(f"Permutation Cramér’s V: observed={V_obs:.3f}, p≈{p_perm:.4f}")
    else:
        print("Permutation skipped — Season×Behaviour not 2×2.")
else:
    print("Permutation skipped — no WS rows.")
```

```
Permutation Cramér’s V: observed=0.058, p≈0.0725
```

Figure 24: Code and Output for Permutation Robustness

6. Limitations (Sabbir)

6.1 Limitation Flowchart

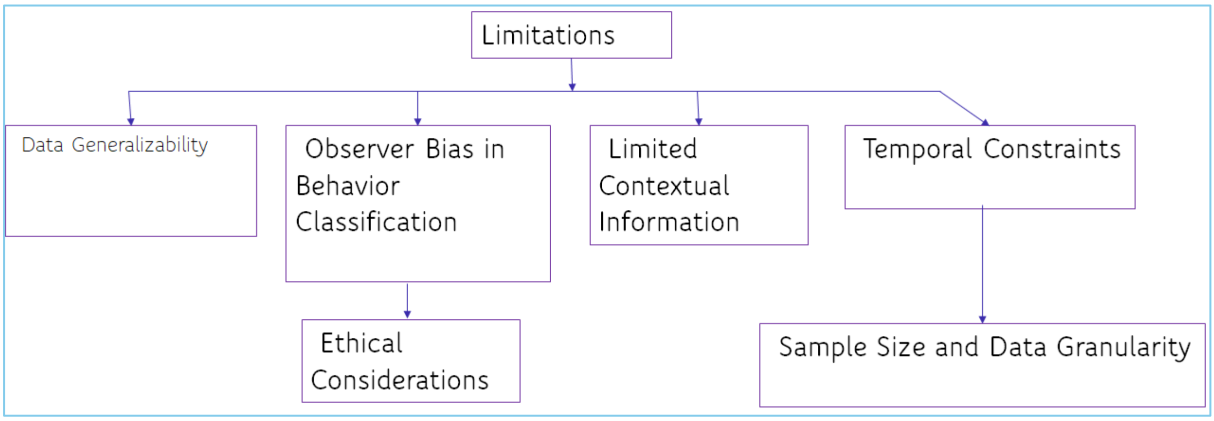

Figure 25: Limitation flowchart format

6.2 Discussion:

This study was not devoid of limitations that had impacts on the scope and the accuracy of its results.

1. **Generalizability of the data:** The research was conducted on a small scale, only on one location and two seasons, which did not allow making any generalizations.
2. **Observer Bias:** There might have been subjective judgement or difference between manual and automated observation when it comes to behavioural classifications.
   Ethical Implications: It is possible that ethical considerations influenced the behaviour of the bats through observation practices with non-invasive data collection being necessary in subsequent research.
3. **Poor Contextual Information:** The lack of ecological parameters, e.g. predators, insects, or moonlight, decreased the richness of behavioural explanation.
4. **Time Limit:** The information was found only over a limited period, which restricted the ability to conduct analysis of seasonal patterns over a long period.
   Sample Size and Granularity: The data coverage was not even across months and thus restricted statistical power and comparison at the fine seasonal scale.

7. Future Scope of Research (Rana)

Although this exploration was insightful in the process of establishing the effects of changes in seasons on the behaviour of bats, there are still a number of opportunities that can be explored further.

1. **Long-term and Multi-season Research:** Future studies need to be conducted over several years and involve all the four seasons to establish the long run behavioural patterns and interannual differences brought about by change in climate or habitat.

2. **Combination of Real-Time Environment Data:** Live meteorological and ecological data (e.g. temperature, humidity, wind speed, insect density, and intensity of moonlight) should be included in the model since it would enable more precise modelling of the effect of environmental signals on foraging behaviour.

3. **State of the art Analytical and Machine Learning models:** Random Forests methods, Neural Networks, and Time-Series Clustering might also be implemented to forecast behaviour patterns, identify abnormalities, as well as find complex and nonlinear relationships between environmental influences and behavioural consequences.

4. **Spatial and Acoustic Tracking:** GPS telemetry, acoustic sensor uses, or infrared imaging would assist in mapping the movements of the bats in real time and also display spatial differences in risk-taking or avoidance behaviours among the various microhabitats.

5. **Comparative Ecological Studies:** Further investigations and experiments with other species and areas would improve the knowledge about how interplays between predators and prey and resource abundance influence the behavioural evolution in the ecosystem.

6. **Ethological and Psychological Intuitions:** Integrating behavioural ecology with cognitive research may be used further to understand how bats perceive risk, retain foraging locations, and make uncertain decisions.

8. Conclusion (Mahdee)

This report examined the changes in the risk-taking and avoidance behaviour of bats over the seasons. The results adequately indicate that seasonality is a strong determinant of behavioural pattern. Cautious energy saving behaviours were exhibited in bats during the winter when food was scarce, and encounters with rat were infrequent. Conversely, the spring led to an abundance of food and increased activity of the rat, which necessitated prolonged and more daring foraging as an adaptive trade-off between hunger and danger.

These trends were supported with the help of statistical analyses. Chi-square tests demonstrated that there was significant association between season, behaviour and reward and the logistic regression showed that timing and activity of rat interact with seasonality to affect behavioural choices. The model was further enhanced by incorporation of environmental variables like temperature, rainfall and moonlight which indicated that bats react to more than one factor.

Ecologically, the findings underscore an ecologically adaptive context-dependent behaviour - bats change their behaviour dynamically with environmental and competitive influences. Although fewer rats are seen during the winter period, their behaviour of caution means they have a sense of danger, and in spring more competition stimulates a thoughtful risk.

Overall, bats apply intelligent, adaptive foraging behaviour that risks and rewards depending on environmental conditions. Further studies are required to expand the period of data acquisition and incorporate real-time ecological parameters to drill down predictive behavioural models. These discoveries provide an in-depth insight into the behavioural adaptation of wildlife to seasonal and ecological change.

Appendix:

1. Reproducibility Statement

All analyses were performed in Google Colab (Python 3.12).

Key libraries: pandas, numpy, matplotlib, scipy, statsmodels, scikit-learn.

Plots and tables were exported to the /figures and /tables directories.

Component	**Output**	**Description**
b_investigation_cleaned.csv	Final analytic dataset	Merged and labelled file
figures/*.png	Visualisations	Figures 4–11
tables/*.csv	Cross-tabulations & OR tables	Used in sections 3–5
Assignment_3_HIT140_Group_10_Sydney.ipynb	Full executable notebook	Reproducibility archive

2. Project GitHub Link:

(Copy the link and paste it to the browser)

https://github.com/Sabbir254/Assignment_3_Inevstigation_B_HIT140_Group_10_Sydney/tree/main