ASSIGNMENT 2

GitHub Repo: https://github.com/amritapathak1/ABM_Frankenstein

Research Question:

How does repeated social rejection influence the emotional trajectory and moral behavior of a stigmatized agent, as modeled through Frankenstein's Creature, and under what social conditions, if any, can early inclusion prevent a shift from peaceful to vengeful behavior?

Introduction:

This project uses agent-based modeling to explore the psychological effects of social rejection, drawing on Mary Shelley's Frankenstein as both narrative inspiration and conceptual framework. In the novel, the Creature begins as an empathetic and peaceful being, but repeated rejection by society leads him toward isolation, resentment, and ultimately violence. The model simulates this emotional transformation using simple agent rules and threshold dynamics.

At its core, the model investigates how sustained social exclusion affects the emotional state of a stigmatized agent. It tracks the Creature's empathy and resentment as they evolve through interactions with different types of human agents—some fearful, some neutral, and some compassionate. These responses are fixed, reflecting social prejudice that precedes behavior. The simulation allows us to examine how early encounters and social environments influence the Creature's long-term moral trajectory.

While inspired by literature, the model speaks to broader questions in social psychology, including how rejection, stigma, and lack of inclusion can shape behavior. It provides a structured way to test how different social conditions might prevent or accelerate emotional breakdown and behavioral deviance. In doing so, it offers insights into both the narrative logic of Frankenstein and the social dynamics that continue to affect marginalized individuals today.

Literature Review:

Over the past two decades, researchers in psychology have explored the deep emotional and behavioral consequences of repeated social rejection. Hutchison, Abrams, and Christian (2007) argue that exclusion threatens three core psychological needs: belonging, self-esteem, and perceived control. When these needs are consistently undermined, individuals experience psychological instability. Many respond with antisocial behavior, while others withdraw entirely. This deviance is not treated as an inherent trait. Instead, their work frames it as a gradual response to ongoing relational harm. Chronic exclusion, they suggest, breaks down an individual's sense of self and disrupts social identity. This process is especially relevant to the emotional transformation seen in Frankenstein's Creature.

Chester and DeWall (2016) expand this discussion by examining the neurobiological effects of rejection. Their research shows that social exclusion activates the same brain regions as physical pain. The emotional wounds caused by rejection are not metaphorical—they are physiologically real. These authors argue that aggression following rejection often serves a self-regulatory purpose. It can restore a sense of control or power in the aftermath of humiliation. This aggression is not purely impulsive. It may be calculated and even emotionally rewarding, especially when directed at those perceived to be responsible for the rejection.

Leary et al. (2003) provide further evidence through their analysis of fifteen school shooting cases. In nearly all of them, the perpetrators had experienced intense social rejection. These experiences included bullying, exclusion, and romantic humiliation. Over time, resentment built up and transformed into a broader sense of injustice. The aggression that followed was rarely directed solely at those responsible. Instead, it became displaced and generalized, often targeting innocent bystanders. This pattern illustrates how rejection can lead to emotionally driven violence, not just retaliatory logic.

Together, these studies offer strong empirical support for modeling how rejection influences behavior. They suggest that emotional deterioration is cumulative. Under sustained social pressure, even empathetic individuals may adopt harmful or aggressive strategies as a form of psychological survival. These findings provide a compelling basis for simulating emotional drift in agent-based models, particularly when exploring the social origins of deviance.

Mary Shelley's Frankenstein vividly dramatizes the psychological toll of persistent social rejection. The Creature begins as a curious, emotionally sensitive, and prosocial figure. He seeks connection, learns language, and tries to emulate human norms. Despite these efforts, he is rejected and attacked in nearly every encounter. This transformation—from empathy to vengeance—is often interpreted as a critique of societal prejudice. While literary scholars have noted this theme for years, more recent work provides a sociological lens to formalize it.

Estefania Alvarez (2023) applies Erving Goffman's (1963) theory of stigma to the Creature's experience. She argues that his visible physical difference represents what Goffman termed an "abomination of the body." This is not merely a difference—it is a discrediting trait that provokes fear and horror. According to this theory, stigma operates before any behavior is judged. The Creature is rejected not for what he does, but for how he appears. His physical form violates social expectations of normalcy. It marks him as dangerous before he even speaks.

Alvarez illustrates how this stigmatization recurs throughout the narrative. The Creature is violently driven away after helping others. He is beaten for saving a child from drowning. He is shot after performing an act of compassion. These repeated incidents form a consistent pattern: society treats the Creature as a monster regardless of his moral intent. His kindness is erased by the surface-level fear his body produces.

Goffman's labeling theory helps explain why this rejection is so enduring. Once someone is marked as deviant, every action is interpreted through that lens. Even neutral or kind behaviors

are seen as threatening or manipulative. Alvarez shows that the Creature internalizes these reactions. Over time, he comes to believe in the image society reflects of him. This internalized stigma corrodes his hope for belonging.

Despite his repeated efforts to integrate through language, observation, and mimicry, he is denied any chance of affiliation. Eventually, this erosion of hope leads him to abandon his pursuit of connection. He turns toward vengeance, not because of innate aggression, but because every alternative has been closed off. This arc echoes Oyserman and Swim's "coping model" of stigma. In their framework, chronically stigmatized individuals often withdraw. When empowerment or recognition is unavailable, they may adopt defensive or even destructive responses. Alvarez's analysis reveals how this model maps directly onto the Creature's narrative, making Frankenstein a case study in the emotional and moral consequences of sustained social exclusion.

This analysis presents the Creature not as inherently monstrous, but as someone shaped—and perhaps even made monstrous—by the society that continually rejects him. His violence is not innate. It is learned, built from a series of failed attempts to connect. The society around him refuses to recognize his humanity. They fear his appearance and project monstrosity onto him. Over time, this treatment erodes his capacity for empathy and fuels resentment. In this light, Frankenstein becomes more than a gothic novel. It becomes a narrative framework for understanding the emotional and moral toll of systemic exclusion.

The arc of the Creature exemplifies dynamics that modern psychological theories also describe. These include the slow erosion of empathy, the rise of accumulated resentment, and the eventual shift toward deviant or aggressive behavior. Importantly, this shift occurs when inclusion is not just denied, but repeatedly withheld despite the individual's efforts to belong.

To simulate this process computationally, the model draws on foundational work in agent-based modeling. These models are designed to capture how small, localized social interactions can lead to large-scale behavioral change. In particular, the work of Deffuant et al. (2002) offers a useful precedent. Their bounded-confidence and relative-agreement models were originally built to explain opinion dynamics in social systems. These models show how agents can gradually become more extreme in their beliefs over time. This happens through repeated interactions with others, especially in cases where agents begin with uncertainty or emotional vulnerability.

Although Deffuant's work focuses on political polarization, the underlying structure is highly relevant to Frankenstein. The Creature begins with prosocial tendencies. He shows curiosity, kindness, and a desire to be accepted. But each negative encounter pushes him further from his original emotional state. Every rejection adds to his resentment. Every act of exclusion weakens his belief in the possibility of human connection. These emotional shifts accumulate gradually. Eventually, they reach a tipping point.

These emotional escalations — empirically grounded and textually illustrated — call for a computational approach capable of modeling gradual, cumulative change. Deffuant's threshold-based dynamics provide exactly this structure.

At this threshold, moderate behavior gives way to deviance. The peaceful Creature becomes vengeful—not because of a sudden event, but because of a slow and measurable drift in his emotional state. This is exactly the kind of transformation that Deffuant's threshold-based dynamics are designed to capture. The model uses this structure to simulate when and how moral drift becomes irreversible. In doing so, it treats violence not as a starting condition, but as an emergent outcome of sustained social failure.

To complement the threshold-driven structure offered by Deffuant et al., the PsychSim framework developed by Pynadath and Marsella (2005) offers a method for modeling agents with rich internal complexity. In this framework, agents are not limited to reactive behavior. Instead, they possess internal mental states, goals, and beliefs. Crucially, they engage in recursive social reasoning. That is, they form beliefs about what other agents believe, and use those beliefs to guide their actions.

This theory-of-mind architecture makes PsychSim particularly useful for modeling characters like the Creature. In Shelley's narrative, the Creature is not a passive figure. He watches, learns, and remembers. He tracks emotional patterns across time. He uses past encounters to shape expectations about future ones. When rejected repeatedly, he does not simply lash out. He recalibrates. He shifts from seeking kindness to anticipating cruelty. PsychSim allows for this kind of adaptation by embedding memory, expectation, and belief updating directly into the agent's behavior.

The framework has already been applied to model complex emotional and social dynamics. Pynadath and Marsella use it to simulate bullying, retaliation, and conflict escalation. In these scenarios, agents begin with peaceful intentions but adopt defensive or aggressive strategies when subjected to sustained threat. These simulations demonstrate that even emotionally aware agents—those capable of empathy and self-regulation—can shift toward aggression under the right conditions. This supports the idea that deviance can be both learned and emotionally rational, particularly when individuals are unable to achieve safety or recognition through peaceful means.

Applied to Frankenstein, this modeling framework provides a strong cognitive foundation for simulating the Creature's psychological evolution. It allows the Creature to be portrayed not just as a reactive agent, but as a reflective and adaptive one. His responses are shaped by what he expects others to do, and those expectations evolve through repeated rejection.

Together, these modeling approaches—Deffuant's threshold-based emotional drift and PsychSim's cognitive realism—form the foundation for this project's agent-based design. Deffuant et al. help define the emotional mechanism by which resentment overtakes empathy, while PsychSim enables the Creature to internalize experience and adjust behavior. Integrating

these two frameworks makes it possible to simulate both what changes in the Creature's emotional state and why those changes occur.

This combination also expands the scope of the project beyond literary analysis. It offers a computational lens through which to examine broader social dynamics, including radicalization, moral breakdown, and the psychological effects of sustained exclusion. In doing so, the model becomes a platform not only for understanding Frankenstein but for exploring contemporary problems where identity, rejection, and deviance intersect.

Social Mechanism Modeled:

The core social mechanism explored by this model is the emotional impact of persistent social rejection. The central idea is that rejection erodes empathy and inflames resentment, especially in visibly stigmatized individuals. When resentment builds over time and eventually outweighs empathy, a critical threshold is crossed. At this point, the individual's moral orientation may shift. The peaceful desire for connection may give way to vengeful behavior. Once this shift occurs, any hostile actions taken by the individual tend to confirm the expectations of those around them. These observers respond with further rejection. The result is a self-reinforcing cycle of exclusion and aggression. The individual becomes what society already feared. The model is designed to simulate this process step by step.

At the micro level, several elements of this emotional drift are embedded in the model. Stigma and prejudice are represented through simplified categories of human agents. Each human is assigned a fixed behavioral type: "fearful," "neutral," or "compassionate." These types determine the likelihood that the human will accept or reject the Creature. Fearful humans always reject. Compassionate ones always accept. Neutral humans may do either, depending on chance. These types remain constant throughout the simulation. This fixed-label design is drawn from Goffman's theory of social deviance. Alvarez (2023) applies Goffman's work to Frankenstein, showing that the Creature is rejected not because of his actions, but because of how he looks. His visible difference is perceived as a threat, regardless of his behavior. The model adopts this assumption as a baseline condition.

The Creature is initialized with high empathy and low resentment. He begins the simulation in a peaceful state. This reflects the way Shelley presents him early in the novel—as a sensitive, observant, and emotionally open character. The model allows the Creature to move from place to place. These places are fixed "landmarks" such as the forest, the village, the cottage, and the market. At each step, he travels to one of these locations and encounters a human agent. This structure mirrors the episodic wandering that characterizes the first half of the novel. The interactions that follow determine the Creature's emotional trajectory.

The goal of the model is to explore how these repeated interactions shape emotional outcomes over time. Each time the Creature is accepted by a human, his empathy increases slightly, and his resentment decreases. Each rejection has the opposite effect. Rejection lowers empathy and raises resentment. These small emotional changes accumulate. Over many steps, the

emotional state of the Creature can shift significantly. This mechanism is grounded in the work of Chester and DeWall (2016). Their research shows that social acceptance can restore emotional balance, while rejection produces psychological pain. They also demonstrate that aggression following rejection can serve as emotional self-regulation. The Creature's emotional rules are designed to reflect this logic. The model captures how small, repeated rejections—especially when no kindness interrupts them—can lead to emotional collapse. It is not one rejection that matters, but the pattern over time. To implement this feedback-driven process, the model uses simple agent roles and spatial routines, detailed next.

When resentment becomes sufficiently high and empathy falls below a critical point, the Creature's behavioral state shifts. In the model, this happens when resentment exceeds a value of 7 and empathy drops below 3. At this moment, the Creature transitions from the "peaceful" state to a "vengeful" one. This shift is not immediate or binary. The model includes an intermediate state called "cautious." This represents growing wariness, emotional guardedness, and the suppression of initial trust. The cautious state acts as a warning zone. It indicates that emotional damage is accumulating, but has not yet tipped into aggression. This multi-stage design reflects the threshold logic described by Deffuant et al. (2002). Their work shows that emotional or ideological drift can be slow and continuous, until a threshold is crossed. After that, behavior can change suddenly and sharply.

The model also captures a feedback loop. Emotional changes do not remain internal. They influence the Creature's posture and outward behavior. As he becomes more cautious or resentful, his demeanor changes. He may approach others less openly. His movements may appear hesitant, tense, or defensive. For fearful human agents, this shift can be misread. Caution is interpreted as a threat. The Creature begins to look dangerous, regardless of his intent. This increases the likelihood of rejection. Repeated rejection then deepens the Creature's alienation. In this way, emotional drift leads to behavioral change. That change invites more exclusion. The process feeds on itself.

This loop resembles a self-fulfilling prophecy. The more the Creature is feared, the more he becomes what people fear. This mirrors the core dynamic of Shelley's novel. In Frankenstein, the Creature does not begin with violence. Violence emerges only after society has exhausted every path to peaceful connection. The model attempts to formalize this arc. It also reflects the kind of stigmatization process Goffman described. Once an individual is marked as deviant, all behavior is interpreted through that lens. Even kindness becomes suspicious.

This feedback structure also parallels real-world patterns of exclusion and aggression. Leary et al. (2003) studied cases of school shootings. They found that repeated rejection and social humiliation were common precursors to violence. The emotional build-up in those cases followed a similar path—accumulated rejection, rising resentment, and eventual aggression. The model draws on this work to simulate how deviance can emerge gradually, not from pathology, but from relational harm.

Although the model is still in development, its current structure suggests several likely patterns. In environments where a large proportion of humans are fearful, such as 80%, the Creature is likely to reach the vengeful state quickly. This might occur within 20 time steps. Early rejection accelerates emotional collapse. On the other hand, if the Creature meets a compassionate agent early in the simulation, this may alter the outcome. A single act of kindness can preserve empathy. It can slow the rise of resentment. In these cases, the Creature may remain peaceful or only become cautious. In mixed societies, outcomes become less predictable. The early sequence of interactions begins to matter more. Some simulations end in peace. Others end in aggression. This reflects the fragility of moral development under stress. The same agent, under different early conditions, may follow dramatically different paths.

Ultimately, the goal of the model is to formalize and simulate a process of socially induced moral drift. In this process, deviance does not originate from internal pathology. It emerges from repeated experiences of exclusion, rejection, and failed social integration. The model treats harmful behavior as the consequence of cumulative relational harm, not as a reflection of innate aggression. The Creature becomes deviant not because he is evil, but because society gives him no viable path toward belonging.

While the model is still under development, it offers a useful framework for experimentation. It allows us to test how different patterns of early inclusion or marginalization might shape long-term emotional and behavioral outcomes. By adjusting the social environment—changing the proportion of fearful versus compassionate agents—we can observe how small changes in social context might alter the Creature's trajectory. These simulations are designed to explore not just what happens, but why it happens. They trace the mechanisms by which deviance can emerge over time.

Although the model takes inspiration from Frankenstein, its implications extend beyond literary analysis. It speaks to broader social concerns. These include the psychological roots of radicalization, the emotional toll of repeated exclusion, and the long-term challenges faced by stigmatized individuals reentering society—such as formerly incarcerated people or those marked by visible difference. In this way, the model bridges literature, psychology, and computational social science. It offers a structured way to explore how society's treatment of outsiders shapes the moral and emotional futures of those it excludes.

Model Structure and Operationalization of Agents

The model simulates a simplified social world based on Frankenstein, where the Creature repeatedly encounters human agents who respond to him with varying degrees of acceptance or rejection. Each agent in the model is designed to represent a specific social role and emotional function. The Creature is the central learning agent. He begins the simulation in a peaceful emotional state. His empathy is set high. His resentment is set low. These emotional attributes change gradually based on interactions with others.

The human agents represent society's range of responses to visible deviance or unfamiliarity. Each human is assigned one of three fixed behavioral types: "fearful," "neutral," or "compassionate." Fearful humans always reject the Creature. Compassionate humans always accept him. Neutral humans respond randomly, sometimes accepting, sometimes rejecting. These types do not change over time. They operationalize how social stigma works as a structural force. Their reactions are not based on the Creature's behavior. They are based on his appearance, as described by Goffman and applied by Alvarez (2023).

The Creature moves between a small number of fixed locations, such as the forest, village, cottage, or market. At each step, he visits one of these landmarks and encounters one or more human agents. The outcome of each interaction affects his emotional state. Acceptance increases empathy and reduces resentment. Rejection decreases empathy and raises resentment. Over time, these changes accumulate. Once resentment becomes high and empathy falls low, the Creature's behavioral state shifts. He may become cautious. He may eventually become vengeful. These state transitions are governed by specific numerical thresholds defined in the model.

This structure creates a feedback loop. The Creature's emotional state influences how he acts. This behavior may then influence how others perceive him. The model is designed to simulate how sustained rejection can produce deviance, even when the individual begins with peaceful intent.

All key features of this model—including agent properties, movement rules, emotional update functions, and threshold transitions—are implemented and documented in the code. The logic behind every design choice is explicitly reflected in the structure of the Python classes and comments in both agent.py and model.py. A reader with basic familiarity with agent-based modeling and Python could reproduce the simulation using only the code and the descriptions provided here. The code serves not only as an implementation but also as a clear expression of the model's conceptual logic.

REFERENCES:

1. Chester, D. S., & DeWall, C. N. (2017). Combating the sting of rejection with the pleasure of revenge: A new look at how emotion shapes aggression. Journal of Personality and Social Psychology, 112(3), 413–430.

2. Alvarez, E. (n.d.). The stigmatization of Frankenstein's monster. Dawson English Journal.

3. Hutchison, P., Abrams, D., & Christian, J. (2007). The social psychology of exclusion. In D. Abrams, J. Christian, & D. Gordon (Eds.), Multidisciplinary handbook of social exclusion research (pp. 27–45). John Wiley & Sons

4. Deffuant, G., Amblard, F., Weisbuch, G., & Faure, T. (2002). How can extremism prevail? A study based on the relative agreement interaction model. Journal of Artificial Societies and Social Simulation, 5(4).

5. Leary, M. R., Kowalski, R. M., Smith, L., & Phillips, S. (2003). Teasing, rejection, and violence: Case studies of the school shootings. Aggressive Behavior, 29, 202–214.

6. Abrams, D., Randsley de Moura, G., Hutchison, P., & Viki, G. T. (2005). When bad becomes good (and vice versa): Why social exclusion is not based on difference. In D. Abrams, J. Marques, & M. A. Hogg (Eds.), The social psychology of inclusion and exclusion (pp. 161–190). Psychology Press

7. Pynadath, D. V., & Marsella, S. C. (2005). PsychSim: Modeling theory of mind with decision-theoretic agents. In Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI) (pp. 1181–1186).

8. Perry, G. L. W., & O'Sullivan, D. (2017). Identifying narrative descriptions in agent-based models representing past human–environment interactions. Journal of Archaeological Method and Theory, 25, 795–817.