Here's a structured **table format** for the **GenetiQStride™ dataset**:

GenetiQStride™ Dataset Structure

Column Name	Data Type	Description
Horse_ID	String	Unique identifier for each horse (e.g., H00001).
Breed	Categorical	Horse breed (**Thoroughbred, Arabian, Standardbred, etc.**).
Age	Integer	Age of the horse (3-7 years, as racehorses peak in this range).
Sex	Categorical	Male (**M**) or Female (**F**).
MSTN_Gene	Categorical	Myostatin gene variant: **C/C (fast-twitch), C/T (mixed), T/T (endurance)**.
PPARδ_Gene	Categorical	Endurance/metabolism gene: **T/T (high), C/T (moderate), C/C (low efficiency)**.
COL1A1_Risk	Categorical	Collagen-related injury risk: **Low, Medium, High**.
ACTN3_Type	Categorical	Sprint vs. endurance muscle fibers: **Fast, Mixed, Slow**.
Race_Wins	Integer	Total races won (0-15).
Avg_Speed (km/h)	Float	Average race speed (48-70 km/h).
Distance_Pref	Categorical	Preferred racing distance: **Sprint, Mid, Long**.
Injury_History	Categorical	Whether the horse had past injuries (**Yes/No**).
Sire_ID	String	Unique ID of the horse's father (e.g., S001).
Dam_ID	String	Unique ID of the horse's mother (e.g., D001).
Champion_Lineage	Categorical	If the horse comes from a championship bloodline (**Yes/No**).
Inbreeding_Coeff	Float	Genetic inbreeding coefficient (0.01 - 0.15).
Race_Potential	Float	Predicted race performance score (0-100).

Column Name	Data Type	Description
Injury_Risk	Float	Predicted likelihood of injury (0-100).

This **structured table** ensures the dataset is well-organized, machine-learning-ready, and biologically accurate. Let me know if you'd like any modifications! 🚀