User Name: Aaryan
User ID: 37275
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

All Categorical Variables:

Embarked

Sex

Survived

Ticket

Name

Cabin

Independent Continuous Variables:

PassengerId

Pclass

Age

SibSp

Parch

Fare

User Name: Aaryan
User ID: 89790
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

Kolmogorov-Smirnov Test Results:

Variable	p-value	KS-Statistic	Inference	Result
PassengerId	1.5682775992087616e-48	0.2684776655443322	Does not match the real data distribution significantly.	Fail
Pclass	6.040481159102964e-12	0.13197575757575758	Does not match the real data distribution significantly.	Fail
Age	3.457973660150079e-39	0.2413117845117845	Does not match the real data distribution significantly.	Fail
SibSp	2.5786292703585823e-05	0.08597934904601572	Does not match the real data distribution significantly.	Fail
Parch	7.984215474120097e-05	0.08154276094276094	Does not match the real data distribution significantly.	Fail
Fare	1.8106059226207038e-157	0.47710864197530867	Does not match the real data distribution significantly.	Fail

User Name: Aaryan
User ID: 30566
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

T-Test Results:

Variable	p-value	T-Statistic	Inference	Result
PassengerId	0.0002773089345938805	3.637819205714891	Significantly different means from real data.	Fail
Pclass	1.0331804101507008e-09	6.114126846359151	Significantly different means from real data.	Fail
Age	nan	nan	Significantly different means from real data.	Fail
SibSp	2.131506513712287e-06	-4.745266668286333	Significantly different means from real data.	Fail
Parch	0.011809978805879971	-2.51855653961849	Significantly different means from real data.	Fail
Fare	1.7469815171730086e-12	7.068543946945586	Significantly different means from real data.	Fail

User Name: Aaryan
User ID: 18137
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

Mann-Whitney U Test Results:

Variable	p-value	MWU-Statistic	Inference	Result
PassengerId	3.0483860501763937e-06	2444126.0	Significantly different distributions from real data.	Fail
Pclass	5.018352759625736e-09	2477037.5	Significantly different distributions from real data.	Fail
Age	nan	nan	Significantly different distributions from real data.	Fail
SibSp	3.7416323677997957e-06	2040726.5	Significantly different distributions from real data.	Fail
Parch	4.279772617887337e-05	2072229.0	Significantly different distributions from real data.	Fail
Fare	5.065045944198622e-73	3058671.5	Significantly different distributions from real data.	Fail

User Name: Aaryan
User ID: 95893
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

ANOVA Results:

Variable	p-value	F-Statistic	Inference	Result
PassengerId	3.0482262645543825e-06	13.233728573468158	Significantly different means from real data.	Fail
Pclass	5.0179995192130826e-09	37.38254709336976	Significantly different means from real data.	Fail
Age	6.782911415772228e-11	8.305689333131532	Significantly different means from real data.	Fail
SibSp	3.74140884657156e-06	22.517555753149274	Significantly different means from real data.	Fail
Parch	4.2795293916982664e-05	6.3431270432550635	Significantly different means from real data.	Fail
Fare	5.0640476230184647e-73	49.96431352990109	Significantly different means from real data.	Fail

User Name: Aaryan
User ID: 36762
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

Kruskal-Wallis Test Results:

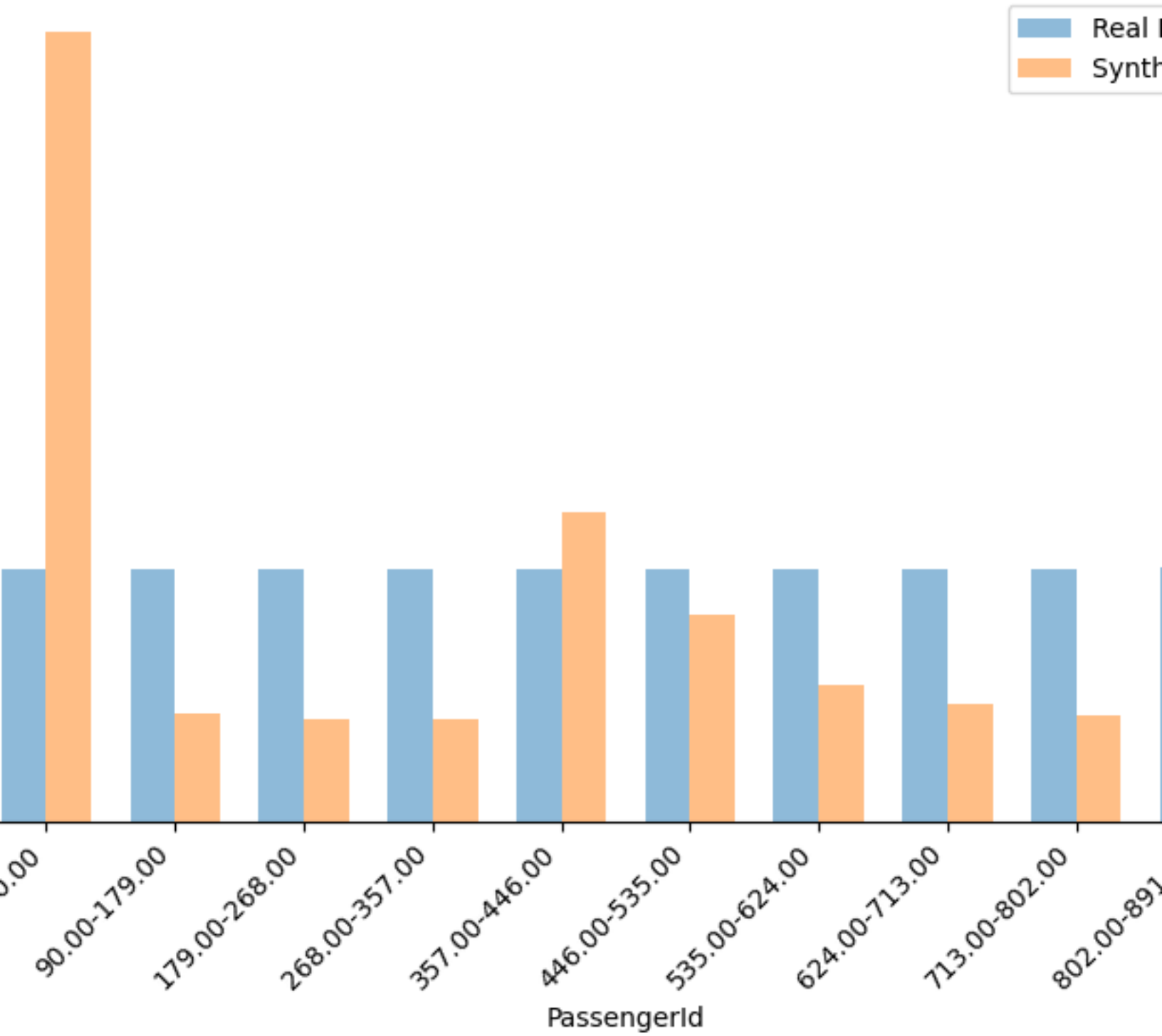
Variable	p-value	H-Statistic	Inference	Result
PassengerId	3.0482262645543825e-06	21.785959811144902	Significantly different distributions from real data.	Fail
Pclass	5.0179995192130826e-09	34.18242925243185	Significantly different distributions from real data.	Fail
Age	6.782911415772228e-11	42.58059080090582	Significantly different distributions from real data.	Fail
SibSp	3.74140884657156e-06	21.392968094764136	Significantly different distributions from real data.	Fail
Parch	4.2795293916982664e-05	16.74320318265976	Significantly different distributions from real data.	Fail
Fare	5.0640476230184647e-73	326.68643198538064	Significantly different distributions from real data.	Fail

User Name: Aaryan
User ID: 78878
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

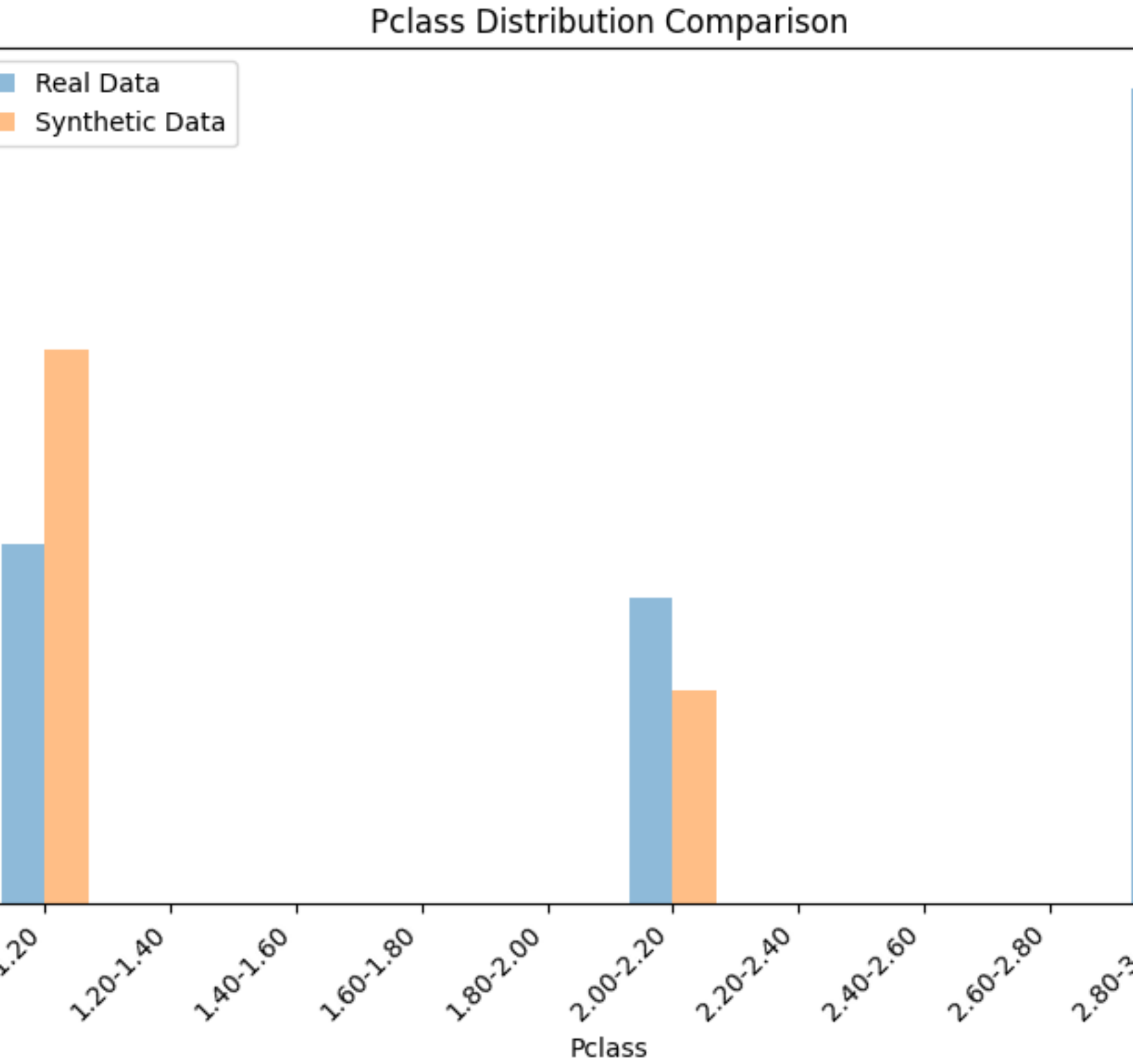
Chi-Square Test Results:

Variable	p-value	Chi2 Statistic	Inference	Result
Embarked	0.17075230342654243	6.406891883591322	Sig. association between the two variables (similar)	Pass
Sex	0.8925041413931792	0.018261808899252534	Sig. association between the two variables (similar)	Pass
Survived	0.5068543903161106	0.44055564653369533	Sig. association between the two variables (similar)	Pass
Ticket	3.5435223047565303e-07	104255.5319669959	No association between the two variables.	Fail
Name	0.31625643680525556	361746.0000000001	Sig. association between the two variables (similar)	Pass
Cabin	0.16315724393042344	1821.25	Sig. association between the two variables (similar)	Pass

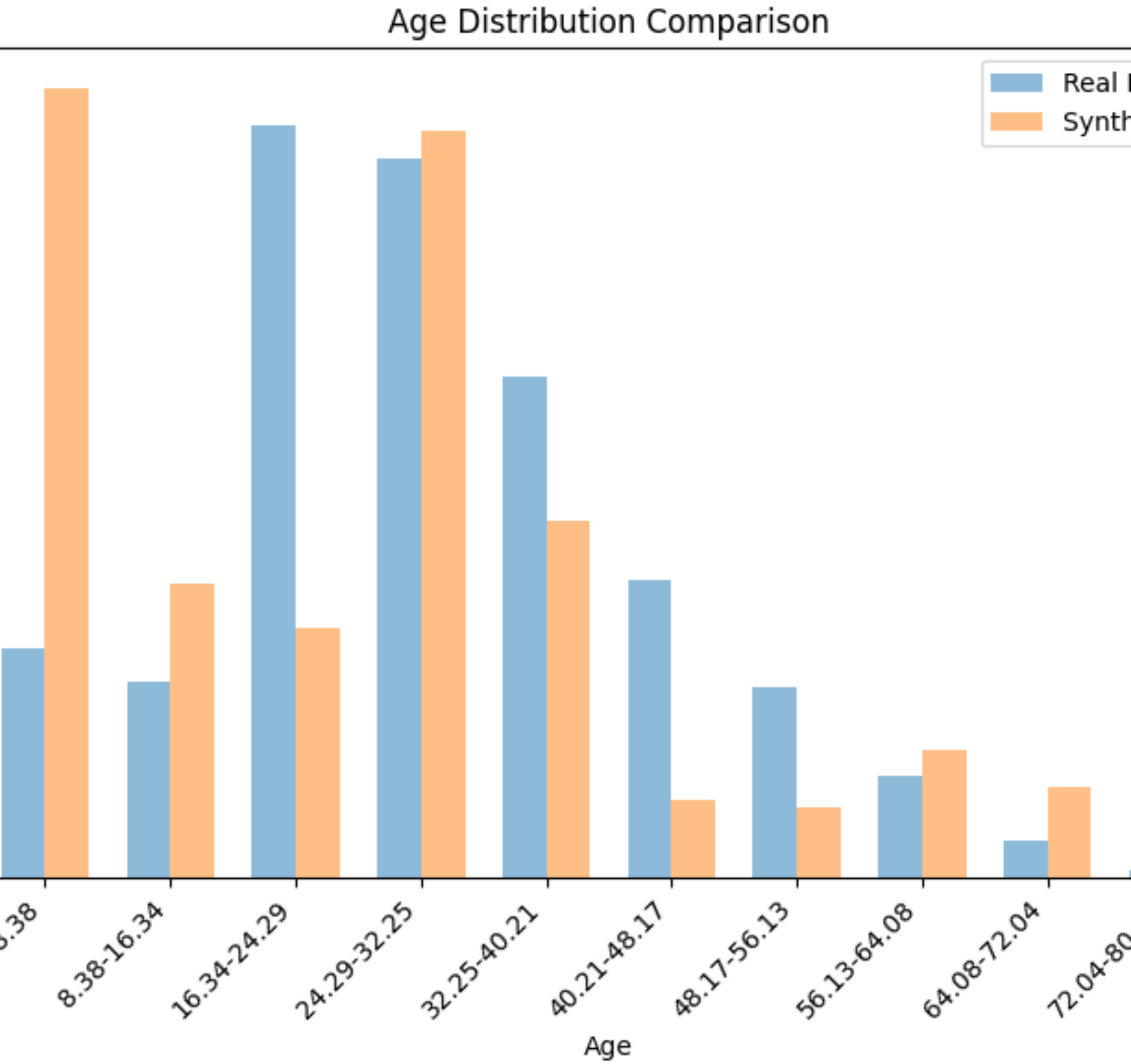
User Name: Aaryan
User ID: 65573
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41


PassengerId Distribution Comparison
Real
Synth
90.00-179.00
179.00-268.00
268.00-357.00
357.00-446.00
446.00-535.00
535.00-624.00
624.00-713.00
713.00-802.00
PassengerId

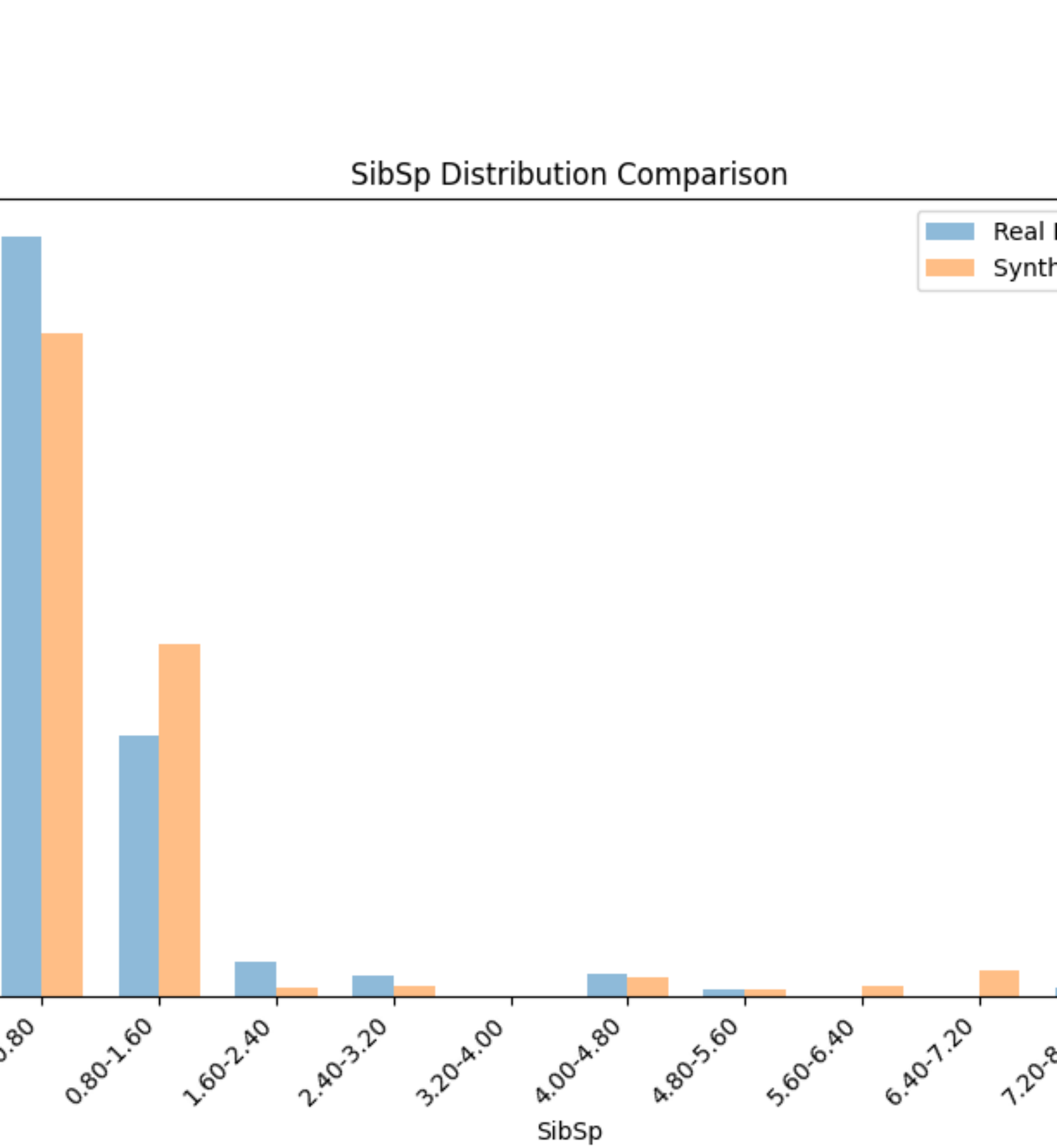
User Name: Aaryan
User ID: 37871
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41


Pclass Distribution Comparison
Real Data
Synthetic Data
1.20-1.40
1.40-1.60
1.60-1.80
1.80-2.00
2.00-2.20
2.20-2.40
2.40-2.60
2.60-2.80
Pclass

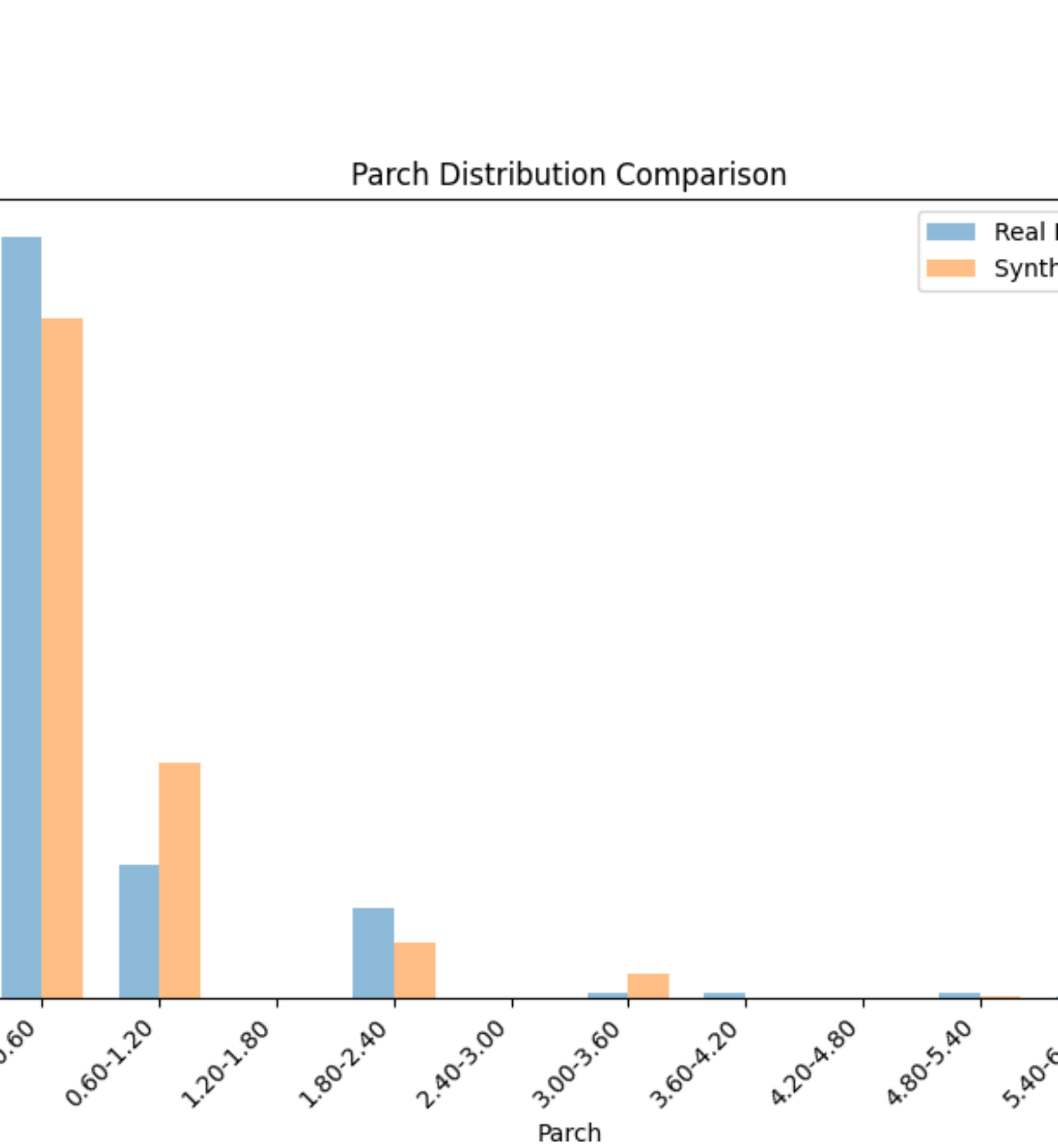
User Name: Aaryan
User ID: 24971
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41


Age Distribution Comparison
Real
Synth
8.38-16.34
16.34-24.29
24.29-32.25
32.25-40.21
40.21-48.17
48.17-56.13
56.13-64.08
64.08-72.04
Age

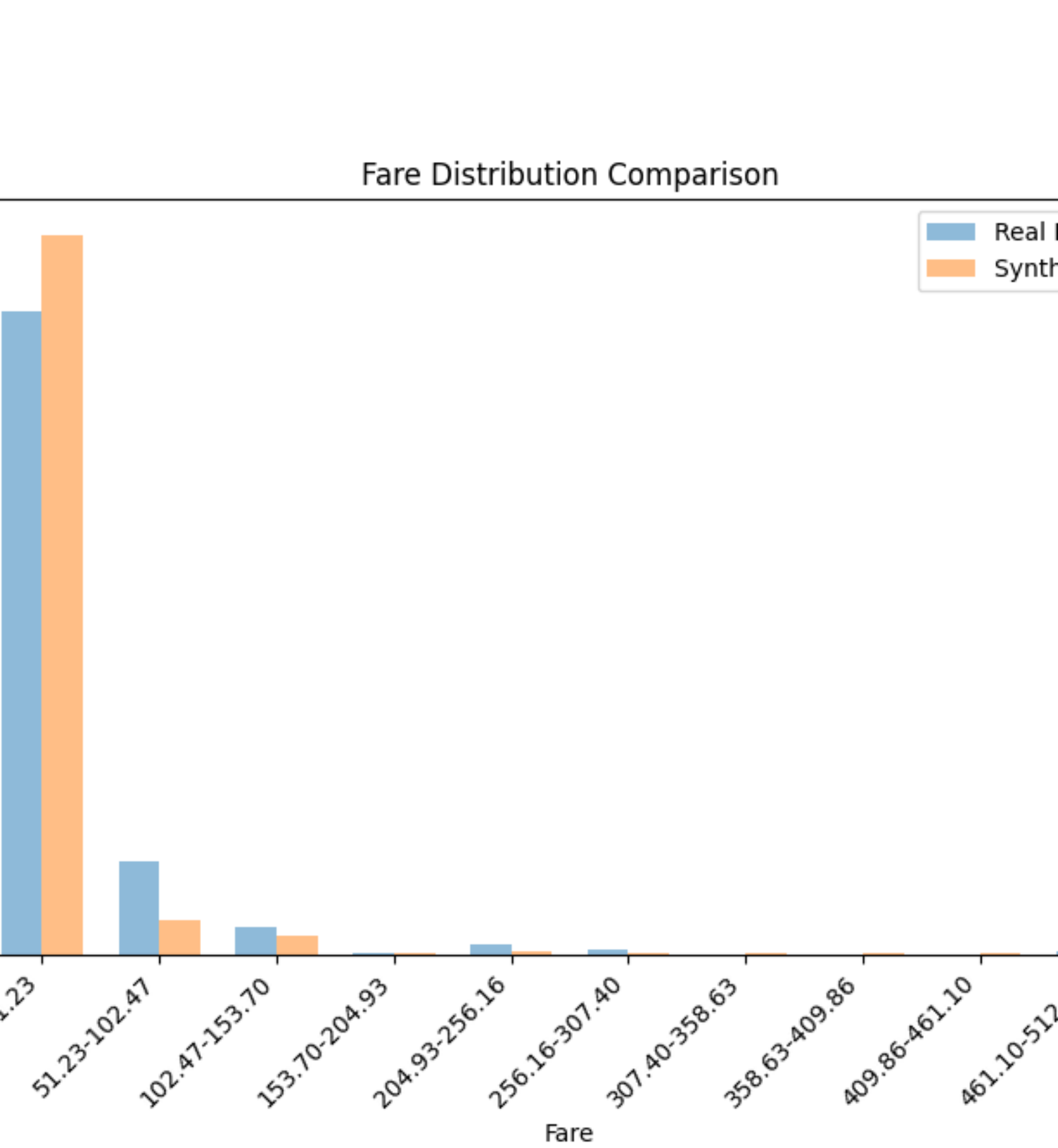
User Name: Aaryan
User ID: 18013
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41


SibSp Distribution Comparison
Real
Synth
0.80
0.80-1.60
1.60-2.40
2.40-3.20
3.20-4.00
4.00-4.80
4.80-5.60
5.60-6.40
6.40-7.20
7.20-8
SibSp

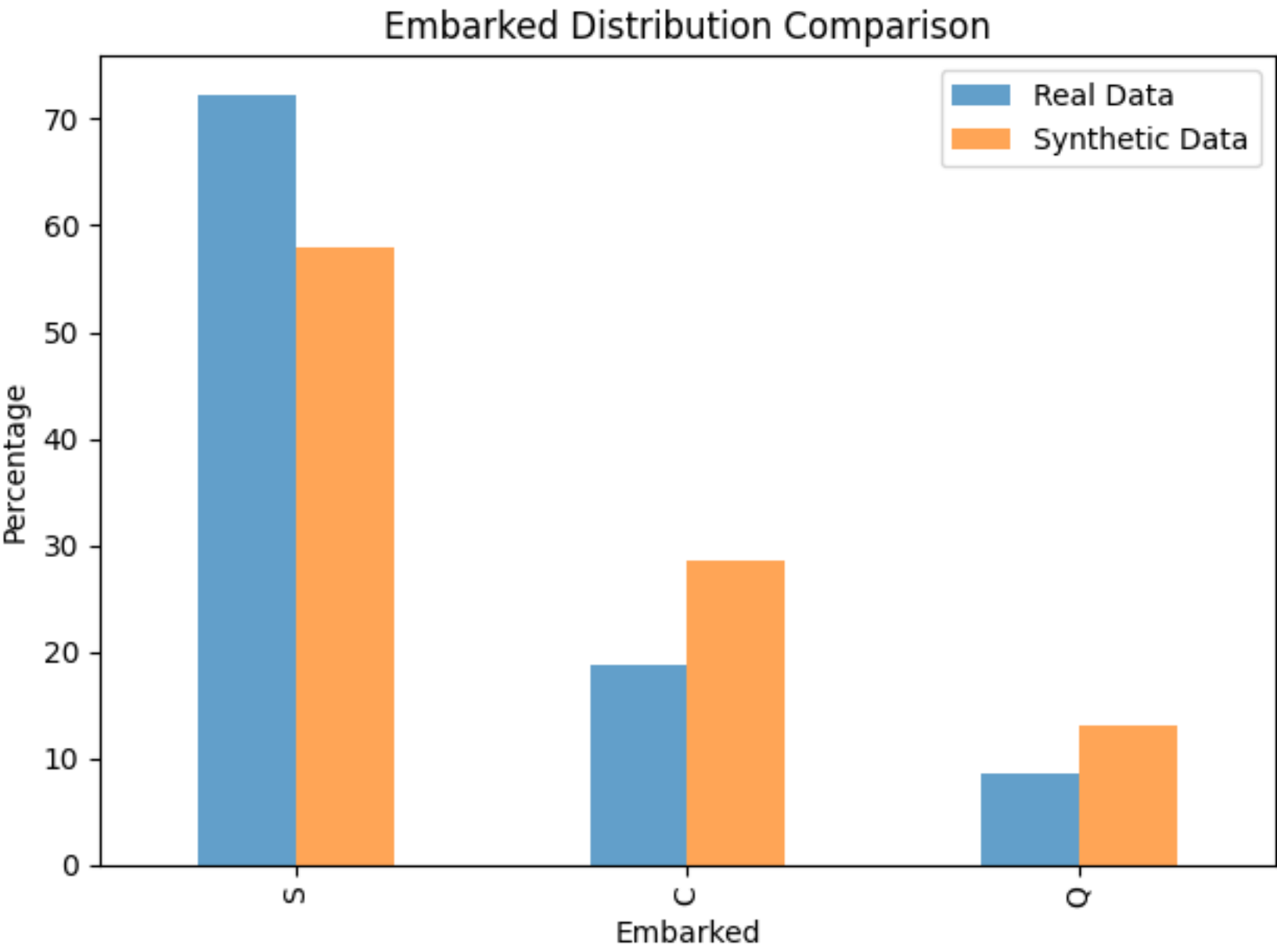
User Name: Aaryan
User ID: 97196
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41


Parch Distribution Comparison
Real
Synth
0.60
0.60-1.20
1.20-1.80
1.80-2.40
2.40-3.00
3.00-3.60
3.60-4.20
4.20-4.80
4.80-5.40
5.40-6
Parch

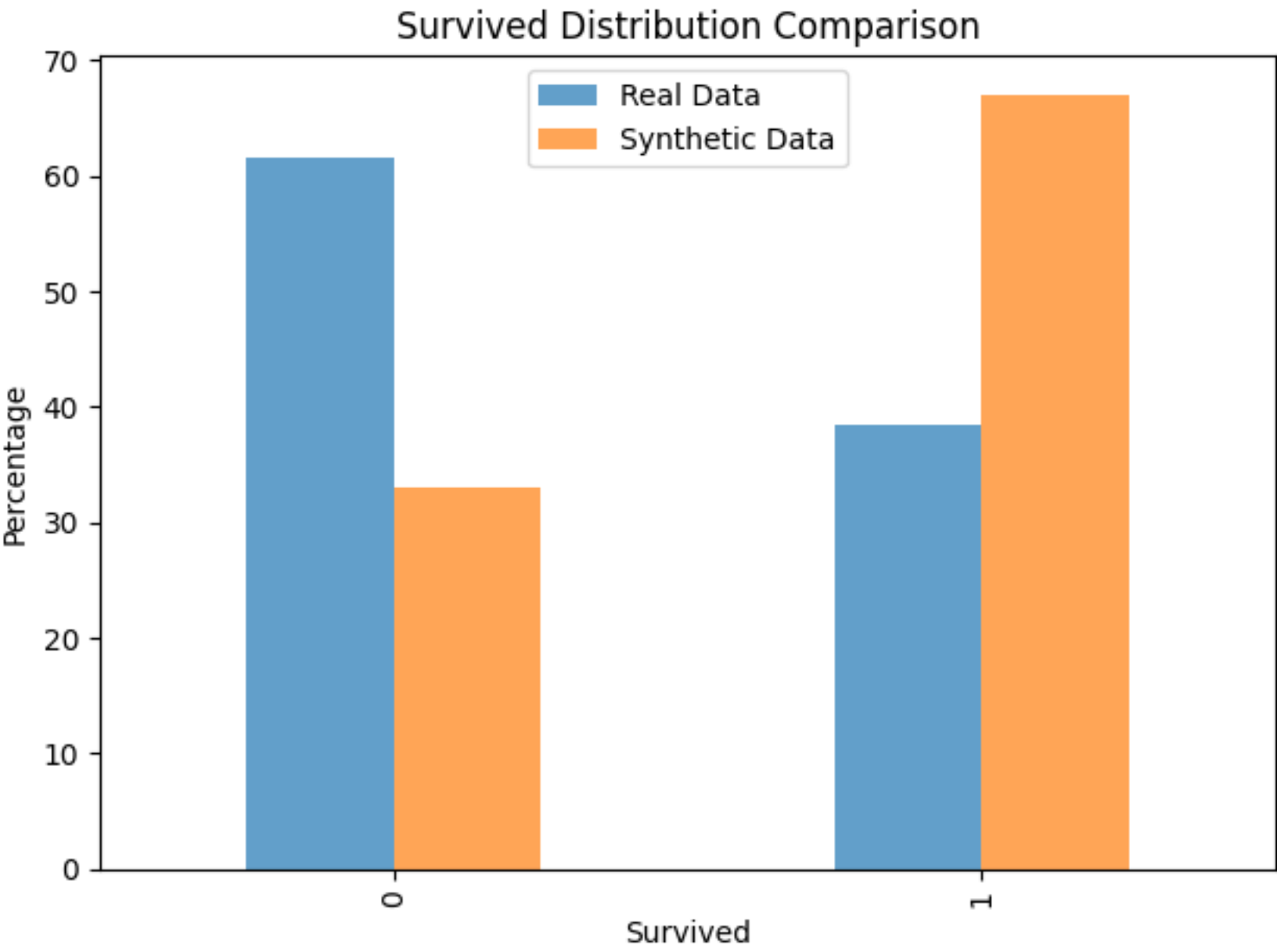
User Name: Aaryan
User ID: 38215
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41


Fare Distribution Comparison
Real
Synth
1.23
51.23-102.47
102.47-153.70
153.70-204.93
204.93-256.16
256.16-307.40
307.40-358.63
358.63-409.86
409.86-461.10
461.10-512
Fare

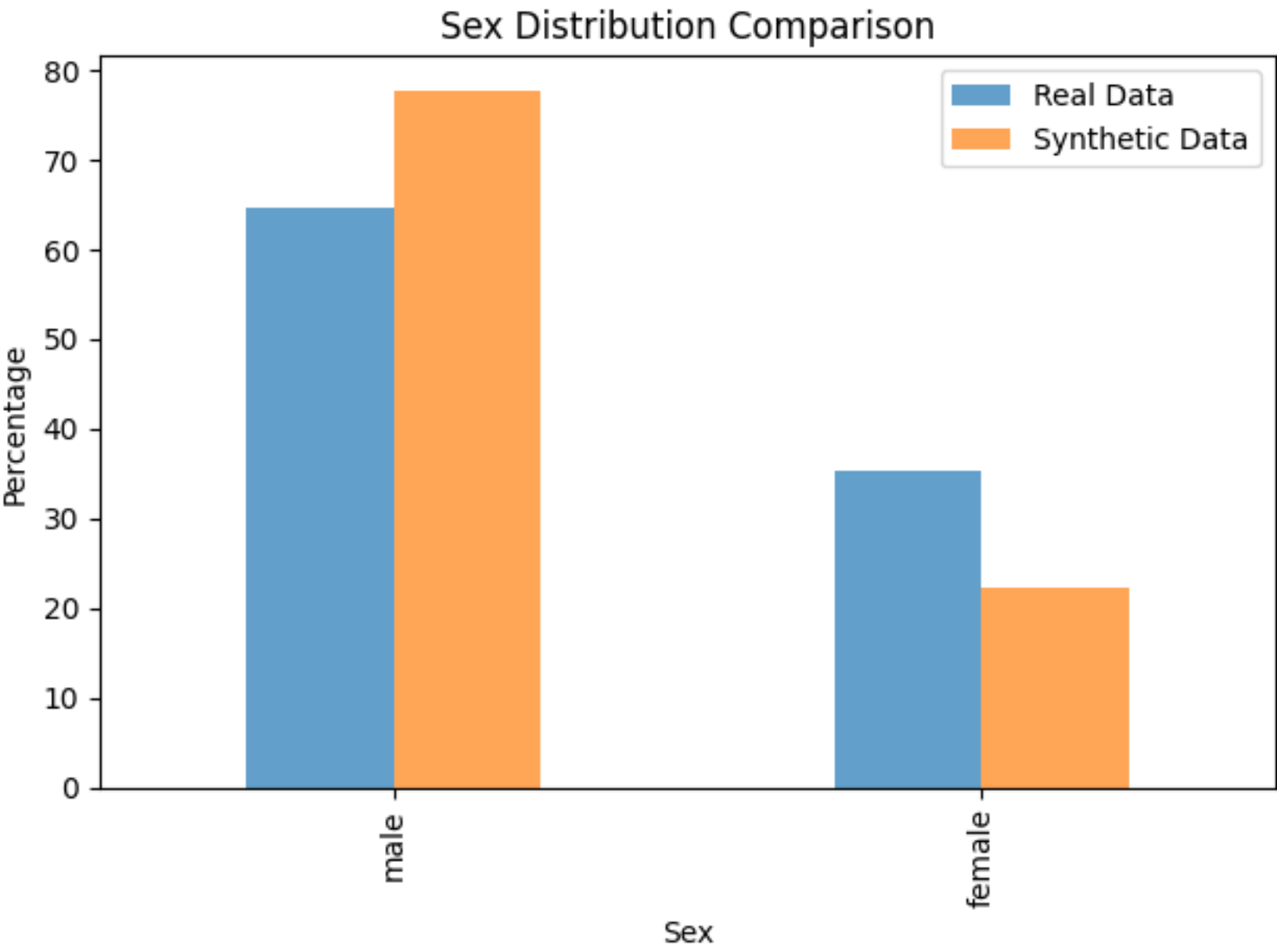
User Name: Aaryan
User ID: 16803
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41



User Name: Aaryan
User ID: 27528
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

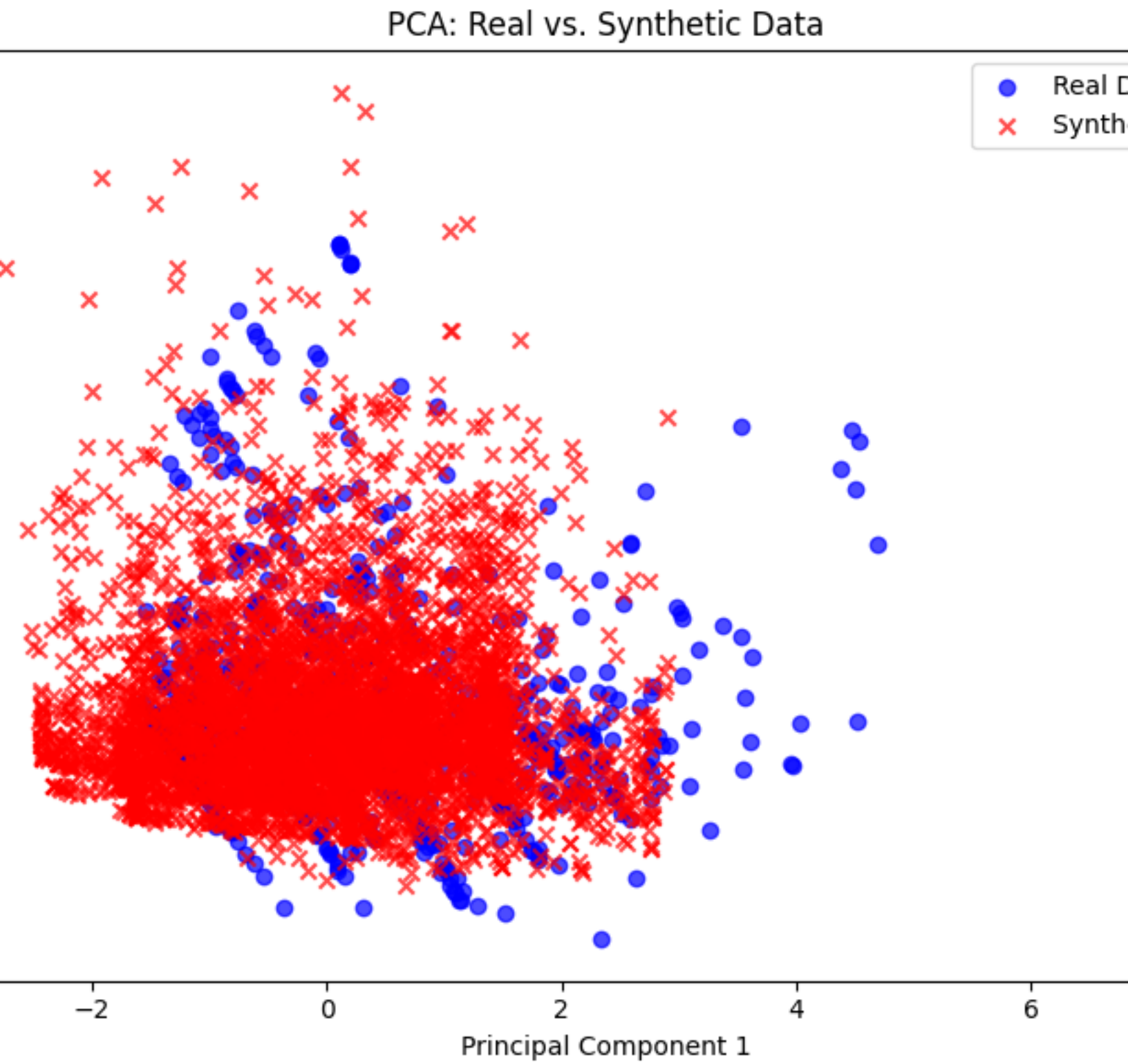


User Name: Aaryan
User ID: 84681
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41



User Name: Aaryan
User ID: 54665
Tool Used: Gretel
Execution Start Time: 2023-08-06 11:54:41
Execution End Time: 2023-08-06 11:54:41

PCA: Real vs. Synthetic Data


PCA: Real vs. Synthetic Data
Real D
Synth
−2
0
2
4
6
Principal Component 1