Resource Allocation and Optimization
Volume 1 · Number 1 · 2026

Spend the Observation Budget on the Target: Feature Allocation for Hidden Digit Blocks

Brandon Franklin[1] · Stephen Young[2] · Ryan King[3]

[1] Belgium · [2] Austria · [3] Italy

Original publication: October 31, 2026 | Article 649

Abstract

Large variation in an observed feature does not necessarily make it useful for predicting a missing region. We compared two training-defined observation menus for four concealed Optdigits features. With sixteen visible features, target-directed greedy selection gave test MSE 0.07122, versus 0.08336 for marginal-variance selection. The advantage persisted with thirty-two observations. Both menus were fixed before the 1,797 test images were evaluated. The result supports task-specific allocation within this numerical setting, while leaving physical acquisition cost and the global optimality of the selected menu unresolved.

Train-only allocation experiment on public block-count features. Observation count is the resource unit; acquisition time, physical scanning, and a specialized inpainting model are not evaluated.

Spend the Observation Budget on the Target: Feature Allocation for Hidden Digit Blocks

> Train-only allocation experiment on public block-count features. Observation count is the resource unit; acquisition time, physical scanning, and a specialized inpainting model are not evaluated.

Abstract

Large variation in an observed feature does not necessarily make it useful for predicting a missing region. We compared two training-defined observation menus for four concealed Optdigits features. With sixteen visible features, target-directed greedy selection gave test MSE 0.07122, versus 0.08336 for marginal-variance selection. The advantage persisted with thirty-two observations. Both menus were fixed before the 1,797 test images were evaluated. The result supports task-specific allocation within this numerical setting, while leaving physical acquisition cost and the global optimality of the selected menu unresolved.

1. Useful observations are defined by a prediction task

A limited observation budget forces a distinction between information that varies and information that helps. A high-variance feature can reveal much about the variability of the observed region while contributing little to a particular missing target. Conversely, a less variable feature may have a strong predictive relationship with the target. Selecting observations solely because they are individually variable can therefore spend the budget on the wrong part of the reconstruction problem.

We ask whether a training-defined, target-directed observation menu improves hidden-feature completion over a menu based on marginal variance at the same feature count. The empirical setting uses sixty candidate visible block counts to predict four central block counts in handwritten digits. Two budgets, sixteen and thirty-two observations, allow us to examine whether an allocation advantage persists when the menu expands. The objective is completion of a specific target, rather than general compression of the visible image.

The application motivation comes from restoration of structured symbols. Meng et al., 2025 develop a zero-shot inpainting model for oracle-bone inscription rubbings. We do not implement that model or treat digit blocks as an archaeological dataset. Instead, the present experiment studies an elementary resource question that can arise before sophisticated restoration: which available context should a fixed predictor receive when the observation menu is constrained?

Our proposition is that a menu selected for the hidden-target objective can outperform a menu selected for high individual feature variance. This is an empirical proposition, not a claim that greedy selection is always optimal. Dependencies among visible features, regularization, finite training data, and test-population differences can all alter the result. The analysis evaluates the specified procedures on the same official held partition.

2. Allocation and reconstruction are related but different choices

Sensor-selection research explicitly distinguishes observation placement from the estimator that later uses the observations. The Gaussian-process placement work of Krause et al., 2008 studies an information-based objective and algorithmic guarantees in its own setting. The convex-optimization approach of Joshi and Boyd, 2009 treats a related subset-choice problem through a different heuristic. These papers motivate a target-aware allocation question, but our covariance-greedy implementation does not inherit their approximation guarantees or reproduce their objectives.

Regression subset methods provide another comparison. Leaps and bounds Furnival and Wilson, 1974 and optimization-based best subset selection Bertsimas et al., 2016 examine alternatives to unrestricted enumeration of predictor sets. We use a forward-only greedy path rather than claiming the best subset among all possibilities. Ridge regularization Hoerl and Kennard, 1970 stabilizes the selected-feature regression. A superior held result for one greedy path does not prove that its selected menu is the globally best use of the budget.

Low-dimensional representation methods Pearson, 1901 and probabilistic PCA Tipping and Bishop, 1999 can also reduce an observation space. They answer a different operational question: a projected component generally requires access to several original measurements. The current budget counts original block features actually supplied to the completion rule. Replacing them with sixteen arbitrary latent components would not preserve that resource definition and is therefore not an additional policy in this experiment.

Matrix completion with noise Candes and Plan, 2010 and spectral regularization Mazumder et al., 2010 motivate modeling relationships among available and missing entries. Modern imputation systems such as GAIN Yoon et al., 2018, GRAPE You et al., 2020, GRIN Cini et al., 2021, and BRITS Cao et al., 2018 address other missing-data structures and modeling assumptions. None is fitted here. They identify possible extensions, while the present study retains a simple common predictor so that the observation menu remains the main difference between conditions.

3. Official data and a fixed concealed target

Optdigits Alpaydin and Kaynak, 1998 provides 3,823 official training records and 1,797 official test records. Each record contains sixty-four values formed from digit-image block counts. We divide counts by sixteen to work on a zero-to-one scale. The four concealed entries have flattened indices 27, 28, 35, and 36, corresponding to the central two-by-two block of the eight-by-eight feature arrangement. All other sixty features are candidates for observation.

Every allocation rule sees the same training feature matrix, including the four target entries during fitting. This supervision is legitimate for the declared prediction task: the targets are known in training and concealed only at test time. Class labels are not used to select features, fit the regression, or evaluate an allocation objective. The policies are global menus, so every test image receives the same selected positions for a given policy and budget.

We estimate the full training mean and covariance with the covariance divisor equal to the number of training records. This convention matters because the regularization constant is specified on the covariance scale. The ridge value is fixed at 0.01 under that convention. Test data do not contribute to the covariance, standardization, selected indices, regularization choice, or sequence of greedy decisions.

The variance policy orders candidate visible features by their marginal training variance, breaking ties by source feature index. Its first sixteen and thirty-two entries define nested menus. This baseline does not

examine covariance with the hidden targets or redundancy between its selected features. It is an intentionally transparent rule for selecting individually variable observations, rather than a claim about an optimal unsupervised observation strategy.

4. A target-directed greedy path

The target-directed policy starts from an empty menu. At each step, it considers adding each remaining candidate and calculates the regularized ridge fitting objective for the four hidden targets. The objective is the total target covariance trace minus the cross-covariance improvement obtained through the selected-feature covariance plus the fixed ridge diagonal. We choose the candidate giving the smallest value, breaking an exact tie by feature index. The first thirty-two choices form one nested path.

This criterion equals a penalized training objective for the corresponding centered linear regression. It is not an unregularized conditional variance, a mutual-information estimate, or a held-out risk. Its trace sums across four target coordinates. In contrast, the reported test MSE averages over those coordinates. The two quantities therefore differ in their normalization and role; the training objective is not printed as though it were a directly comparable test loss.

After a menu is fixed, its predictor uses the same centered ridge solution under either policy. The intercept is the training target mean, and the coefficient matrix is computed from the selected-visible covariance and its cross-covariance with the hidden entries. Predictions are left unclipped. This avoids adding another policy difference after feature selection, although it permits numerical predictions outside the original feature range.

Figure 1 separates the allocation objective from the shared estimator and the held test outcomes. Its arrows show that all menu construction remains inside the training partition. Each policy is evaluated at an equal observation count. The diagram does not imply that these block measurements have equal real acquisition cost or that a physical scanner can obtain only the chosen blocks without seeing the others.

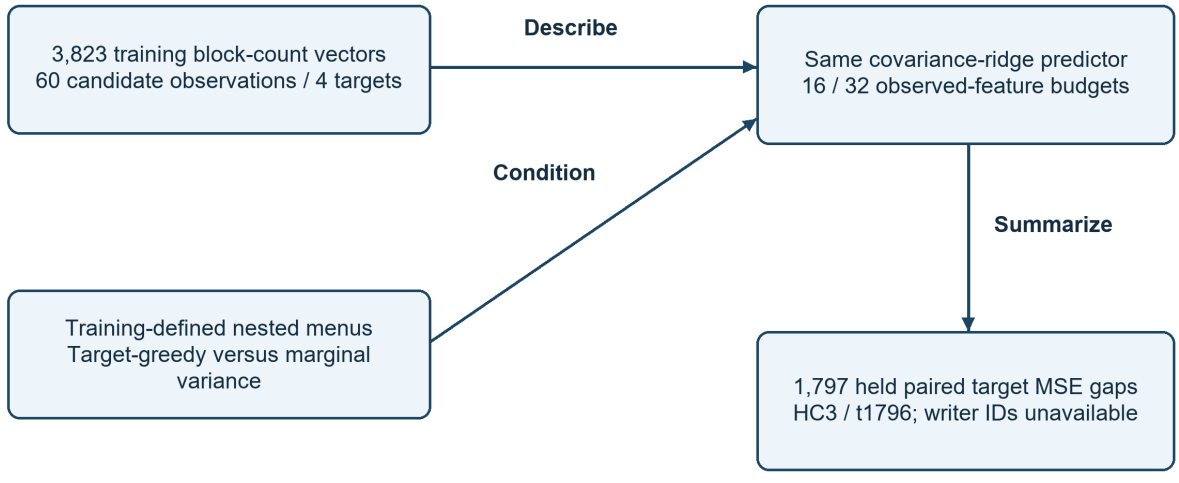


Figure 1. *Marginal variance and target-directed greedy selection supply different nested feature menus. A common covariance-ridge predictor evaluates sixteen- and thirty-two-feature budgets on the same hidden test entries.*

5. Estimation and interpretation of a fixed observation cost

For each test image and budget, we subtract variance-policy MSE from target-greedy MSE. Two intercept-only OLS models summarize the mean paired differences. HC3 uncertainty White, 1980 and t intervals use 1,796 residual degrees of freedom. The observation unit is an image, not an individual hidden block or a candidate feature. The point estimates condition on one fitted covariance and one selected path per policy.

The source documentation describes distinct training and test writers but provides no per-image writer identifier in these feature files. We cannot implement writer-clustered covariance without inventing a grouping. Cluster-inference guidance Cameron and Miller, 2015 therefore informs a stated limitation: the displayed intervals do not account for residual dependence among images from the same writer. They also do not propagate uncertainty from training-data selection or the greedy search.

Observation count is a meaningful numerical constraint, but it is not a measured financial or hardware cost. A digit block can differ in acquisition difficulty or redundancy from another block. Shannon's sampling perspective Shannon, 1949 motivates considering what an observation represents, without supplying a reconstruction theorem for these nonuniform block features. We do not report wall-clock speed, scanner savings, bandwidth, or energy as outcomes.

6. Results at sixteen and thirty-two observations

At the sixteen-feature budget, the target-greedy test MSE is 0.071221 and the variance-policy MSE is 0.083356. Their paired difference is -0.012135, with 95% interval [-0.014226, -0.010044]. At thirty-two features, the corresponding MSE values are 0.067255 and 0.077235. The mean difference remains negative at -0.009980, with interval [-0.011495, -0.008466].

Table 1. Target-greedy minus marginal-variance hidden-feature MSE.

Model	Term	Estimate (SE)	95% CI	N
M1	Mean outcome/contrast	-0.012135 (0.001066)	[-0.014226, -0.010044]	1797
M1: fit	R-squared = 0.0000	-	-	1797
M2	Mean outcome/contrast	-0.009980 (0.000772)	[-0.011495, -0.008466]	1797
M2: fit	R-squared = 0.0000	-	-	1797

Model labels: M1 = Target-greedy minus high-variance allocation at 16 visible features; M2 = Target-greedy minus high-variance allocation at 32 visible features.

Both columns use the same 1,797 official test records, no regression controls, and HC3 t-based uncertainty. The intercept-only R-squared values are effectively zero. The negative coefficients favor target-directed menus under the stated numerical loss. They do not show that every image improves, that additional features always reduce test error, or that the selected subsets are optimal among all possible subsets.

The fitted target-greedy menu at sixteen observations even has lower mean test MSE than the variance menu at thirty-two. This is a descriptive cross-budget comparison from the same experiment, not a complete resource-efficiency frontier. Offline selection cost is omitted, and the policies can have different practical acquisition constraints. The result shows how the chosen allocation criterion can matter as much as the count of observations supplied to this particular predictor.

Unclipped outputs also deserve disclosure. Across the 7,188 hidden-coordinate predictions, the target-greedy menus produce 266 and 290 out-of-range values at the two budgets. The variance menus produce 206 and 240. Thus, the policy with lower mean squared error is not the policy with fewer range violations. We retain the declared loss and report the violations rather than retroactively clipping one method to improve another outcome.

7. Relationship to image restoration and measurement

Context-based restoration Pathak et al., 2016, partial convolutions Liu et al., 2018, and gated convolutions Yu et al., 2018b illustrate models that exploit observed context in different ways. Edge-guided completion Nazeri et al., 2019, denoising representations Vincent et al., 2008, and large-mask restoration Suvorov et al., 2021 add other structures. Our small feature experiment cannot compare those architectures, but it identifies an evaluation variable they can leave implicit: the allocation rule determining which context is available.

Measurement-error analysis Carroll et al., 2006 limits interpreting block-count MSE as visual fidelity. The block representation has already discarded stroke-level structure, and the hidden region is a fixed artificial omission. Influence considerations Cook, 1977 also caution that mean error can reflect particular difficult images. We do not label the remaining variation a semantic failure, measure human readability, or infer the recovery of a specific handwritten stroke.

The numerical work uses NumPy Harris et al., 2020 and SciPy Virtanen et al., 2020 with retained source data, selection logs, and per-image errors. FAIR principles Wilkinson et al., 2016 and dataset documentation Gebru et al., 2021 motivate preserving those relationships. The record makes clear that target entries are available for training selection, which would be an information advantage if an allegedly unsupervised comparator were expected to operate without any target training observations.

8. The cost that an observation count leaves out

The greedy path evaluates sixty candidate additions at its first step, fifty-nine at the next, and eventually twenty-nine at the thirty-second. This is a total of 1,424 candidate-menu objectives, excluding the covariance calculation shared by the policies. The variance baseline only orders the marginal variances. We do not time these operations, but their different structures make it clear that an equal test observation count does not imply equal offline selection effort.

That distinction matters when the menu will be used repeatedly. A more expensive training-stage design could be worthwhile if the same menu is deployed over many future images. Conversely, it might be inappropriate when the training distribution changes frequently and the menu must be rebuilt. Neither deployment frequency nor retraining cost is measured here. The experiment therefore establishes a conditional prediction comparison at equal observed-feature counts, rather than a complete economic optimum.

Regularization conventions also need to travel with a reproduction. The constant 0.01 is added to a covariance matrix divided by the training sample count. A library interface that minimizes an unnormalized sum of squared residuals with a penalty called 0.01 would not fit the same model. Reproducing the analysis requires preserving the objective's normalization or scaling the penalty accordingly. Merely copying the displayed number into another ridge-regression implementation can change the predictor and the selected path.

The observation sets are global rather than image-adaptive. The greedy rule does not inspect a test image and decide which block to request next. Every held image receives the same sixteen or thirty-two source

positions under a given policy. This limits claims about active sensing but also removes a possible information leak: the menu cannot adapt to hidden test entries or to a conveniently recognized test class. Per-image acquisition would require a sequential protocol describing which values become available before each selection.

Finally, the target choice is substantive. Four central block counts emphasize a particular part of digit structure. A menu selected for that target may be poorly suited to a corner gap or a full-image objective. A future allocation study should define its intended missing-region distribution before fitting the menu and should evaluate transfer to independent damage patterns. The present result cannot be carried to a new target simply by retaining the same list of selected indices.

9. Conclusion

For a fixed central hidden target, selecting visible features for a supervised ridge objective yields lower held-feature error than selecting them for marginal variance. The advantage remains at two equal-count budgets. The evidence supports distinguishing observation allocation from observation quantity, while limiting the claim to a fixed numerical feature task, fitted menus, and test partition. Physical acquisition cost and broader restoration quality require their own measurements.

References

1. Liu, Guilin; Reda, Fitsum A.; Shih, Kevin J.; Wang, Ting-Chun; Tao, Andrew; Catanzaro, Bryan. (2018). Image Inpainting for Irregular Holes Using Partial Convolutions. arXiv:1804.07723. https://arxiv.org/abs/1804.07723.
2. Tipping, Michael E.; Bishop, Christopher M.. (1999). Probabilistic Principal Component Analysis. Journal of the Royal Statistical Society Series B: Statistical Methodology. https://doi.org/10.1111/1467-9868.00196.
3. You, Jiaxuan; Ma, Xiaobai; Ding, Daisy Yi; Kochenderfer, Mykel; Leskovec, Jure. (2020). Handling Missing Data with Graph Representation Learning. arXiv:2010.16418. https://arxiv.org/abs/2010.16418.
4. Krause, Andreas; Singh, Ajit; Guestrin, Carlos. (2008). Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies. Journal of Machine Learning Research,9(8),235-284. https://www.jmlr.org/papers/v9/krause08a.html.
5. Alpaydin, E.; Kaynak, C. (1998). Optical Recognition of Handwritten Digits [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C50P49.
6. White, Halbert. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica*, vol. 48, no. 4, 1980, pp. 817-838. https://doi.org/10.2307/1912934.
7. Suvorov, Roman; Logacheva, Elizaveta; Mashikhin, Anton; Remizova, Anastasia; Ashukha, Arsenii; Silvestrov, Aleksei; Kong, Naejin; Goka, Harshith; Park, Kiwoong; Lempitsky, Victor. (2021). Resolution-robust Large Mask Inpainting with Fourier Convolutions. arXiv:2109.07161. https://arxiv.org/abs/2109.07161.
8. Yu, Jiahui; Lin, Zhe; Yang, Jimei; Shen, Xiaohui; Lu, Xin; Huang, Thomas. (2018b). Free-Form Image Inpainting with Gated Convolution. arXiv:1806.03589. https://arxiv.org/abs/1806.03589.
9. Shannon, C.E.. (1949). Communication in the Presence of Noise. Proceedings of the IRE. https://doi.org/10.1109/jrproc.1949.232969.
10. Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S.; Heringa, Jaap; ’t Hoen, Peter A.C; Hooft, Rob; Kuhn, Tobias; Kok, Ruben; Kok, Joost; Lusher, Scott J.; Martone, Maryann E.; Mons, Albert; Packer, Abel L.; Persson, Bengt; Rocca-Serra, Philippe; Roos, Marco; van Schaik, Rene; Sansone, Susanna-Assunta; Schultes, Erik; Sengstag, Thierry; Slater, Ted; Strawn, George; Swertz, Morris A.; Thompson, Mark; van der Lei, Johan; van Mulligen, Erik; Velterop, Jan; Waagmeester, Andra; Wittenburg, Peter; Wolstencroft, Katherine; Zhao, Jun; Mons, Barend. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. https://doi.org/10.1038/sdata.2016.18.
11. Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu. “Measurement Error in Nonlinear Models.” 2006. https://doi.org/10.1201/9781420010138.

12. Furnival, George M.; Wilson, Robert W.. (1974). Regressions by Leaps and Bounds. Technometrics. https://doi.org/10.1080/00401706.1974.10489231.

13. Meng, Zijie; Zeng, Yuanze; Chang, Xiang; Xu, Tianshuo; Chao, Fei; Cao, Xixin; Shang, Changjing; Shen, Qiang. (2025). Orpaint: a zero-shot inpainting model for oracle bone inscription rubbings with visual mamba block. Science China Information Sciences, 68(8), 189102. https://doi.org/10.1007/s11432-024-4493-4.

14. Virtanen, Pauli; Gommers, Ralf; Oliphant, Travis E.; Haberland, Matt; Reddy, Tyler; Cournapeau, David; Burovski, Evgeni; Peterson, Pearu; Weckesser, Warren; Bright, Jonathan; van der Walt, Stéfan J.; Brett, Matthew; Wilson, Joshua; Millman, K. Jarrod; Mayorov, Nikolay; Nelson, Andrew R. J.; Jones, Eric; Kern, Robert; Larson, Eric; Carey, C J; Polat, İlhan; Feng, Yu; Moore, Eric W.; VanderPlas, Jake; Laxalde, Denis; Perktold, Josef; Cimrman, Robert; Henriksen, Ian; Quintero, E. A.; Harris, Charles R.; Archibald, Anne M.; Ribeiro, Antônio H.; Pedregosa, Fabian; van Mulbregt, Paul; SciPy 1.0 Contributors; Vijaykumar, Aditya; Bardelli, Alessandro Pietro; Rothberg, Alex; Hilboll, Andreas; Kloeckner, Andreas; Scopatz, Anthony; Lee, Antony; Rokem, Ariel; Woods, C. Nathan; Fulton, Chad; Masson, Charles; Häggström, Christian; Fitzgerald, Clark; Nicholson, David A.; Hagen, David R.; Pasechnik, Dmitrii V.; Olivetti, Emanuele; Martin, Eric; Wieser, Eric; Silva, Fabrice; Lenders, Felix; Wilhelm, Florian; Young, G.; Price, Gavin A.; Ingold, Gert-Ludwig; Allen, Gregory E.; Lee, Gregory R.; Audren, Hervé; Probst, Irvin; Dietrich, Jörg P.; Silterra, Jacob; Webber, James T; Slavič, Janko; Nothman, Joel; Buchner, Johannes; Kulick, Johannes; Schönberger, Johannes L.; de Miranda Cardoso, José Vinícius; Reimer, Joscha; Harrington, Joseph; Rodríguez, Juan Luis Cano; Nunez-Iglesias, Juan; Kuczynski, Justin; Tritz, Kevin; Thoma, Martin; Newville, Matthew; Kümmerer, Matthias; Bolingbroke, Maximilian; Tartre, Michael; Pak, Mikhail; Smith, Nathaniel J.; Nowaczyk, Nikolai; Shebanov, Nikolay; Pavlyk, Oleksandr; Brodtkorb, Per A.; Lee, Perry; McGibbon, Robert T.; Feldbauer, Roman; Lewis, Sam; Tygier, Sam; Sievert, Scott; Vigna, Sebastiano; Peterson, Stefan; More, Surhud; Pudlik, Tadeusz; Oshima, Takuya; Pingel, Thomas J.; Robitaille, Thomas P.; Spura, Thomas; Jones, Thouis R.; Cera, Tim; Leslie, Tim; Zito, Tiziano; Krauss, Tom; Upadhyay, Utkarsh; Halchenko, Yaroslav O.; Vázquez-Baeza, Yoshiki. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. https://doi.org/10.1038/s41592-019-0686-2.

15. Bertsimas, Dimitris; King, Angela; Mazumder, Rahul. (2016). Best subset selection via a modern optimization lens. The Annals of Statistics. https://doi.org/10.1214/15-aos1388.

16. Hoerl, Arthur E.; Kennard, Robert W.. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics. https://doi.org/10.1080/00401706.1970.10488634.

17. Candes, Emmanuel J; Plan, Yaniv. (2010). Matrix Completion With Noise. Proceedings of the IEEE. https://doi.org/10.1109/jproc.2009.2035722.

18. Harris, Charles R.; Millman, K. Jarrod; van der Walt, Stéfan J.; Gommers, Ralf; Virtanen, Pauli; Cournapeau, David; Wieser, Eric; Taylor, Julian; Berg, Sebastian; Smith, Nathaniel J.; Kern, Robert; Picus, Matti; Hoyer, Stephan; van Kerkwijk, Marten H.; Brett, Matthew; Haldane, Allan; del Río, Jaime Fernández; Wiebe, Mark; Peterson, Pearu; Gérard-Marchant, Pierre; Sheppard, Kevin; Reddy, Tyler; Weckesser, Warren; Abbasi, Hameer; Gohlke, Christoph; Oliphant, Travis E.. (2020). Array programming with NumPy. Nature. https://doi.org/10.1038/s41586-020-2649-2.

19. Cao, Wei; Wang, Dong; Li, Jian; Zhou, Hao; Li, Lei; Li, Yitan. (2018). BRITS: Bidirectional Recurrent Imputation for Time Series. arXiv:1805.10572. https://arxiv.org/abs/1805.10572.

20. Cook, R. Dennis. (1977). Detection of Influential Observation in Linear Regression. Technometrics. https://doi.org/10.1080/00401706.1977.10489493.

21. Joshi, S.; Boyd, S.. (2009). Sensor Selection via Convex Optimization. IEEE Transactions on Signal Processing. https://doi.org/10.1109/tsp.2008.2007095.

22. Vincent, Pascal; Larochelle, Hugo; Bengio, Yoshua; Manzagol, Pierre-Antoine. (2008). Extracting and composing robust features with denoising autoencoders. Proceedings of the 25th international conference on Machine learning - ICML '08. https://doi.org/10.1145/1390156.1390294.

23. Cameron, A. Colin, and Douglas L. Miller. “A Practitioner’s Guide to Cluster-Robust Inference.” *Journal of Human Resources*, vol. 50, no. 2, 2015, pp. 317-372. https://doi.org/10.3368/jhr.50.2.317.

24. Pathak, Deepak; Krahenbuhl, Philipp; Donahue, Jeff; Darrell, Trevor; Efros, Alexei A.. (2016). Context Encoders: Feature Learning by Inpainting. arXiv:1604.07379. https://arxiv.org/abs/1604.07379.

25. Cini, Andrea; Marisca, Ivan; Alippi, Cesare. (2021). Filling the G_ap_s: Multivariate Time Series Imputation by Graph Neural Networks. arXiv:2108.00298. https://arxiv.org/abs/2108.00298.

26. Nazeri, Kamyar; Ng, Eric; Joseph, Tony; Qureshi, Faisal Z.; Ebrahimi, Mehran. (2019). EdgeConnect: Generative Image Inpainting with Adversarial Edge Learning. arXiv:1901.00212. https://arxiv.org/abs/1901.00212.

27. Mazumder, Rahul; Hastie, Trevor; Tibshirani, Robert. (2010). Spectral Regularization Algorithms for Learning Large Incomplete Matrices. Journal of Machine Learning Research 11, 2287-2322. https://jmlr.org/papers/v11/mazumder10a.html.

28. Yoon, Jinsung; Jordon, James; van der Schaar, Mihaela. (2018). GAIN: Missing Data Imputation using Generative Adversarial Nets. arXiv:1806.02920. https://arxiv.org/abs/1806.02920.

29. Pearson, Karl. (1901). LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. https://doi.org/10.1080/14786440109462720.

30. Gebru, Timnit, et al. "Datasheets for datasets" *Communications of the ACM*, 2021. https://doi.org/10.1145/3458723.