TEST-DRIVEN DATA ANALYSIS

Nicholas J. Radcliffe
Stochastic Solutions Limited
& Department of Mathematics, University of Edinburgh

TEST-DRIVEN DATA ANALYSIS

or

Why should anyone believe your analytical results?

or

Data analysis as if the answers actually mattered

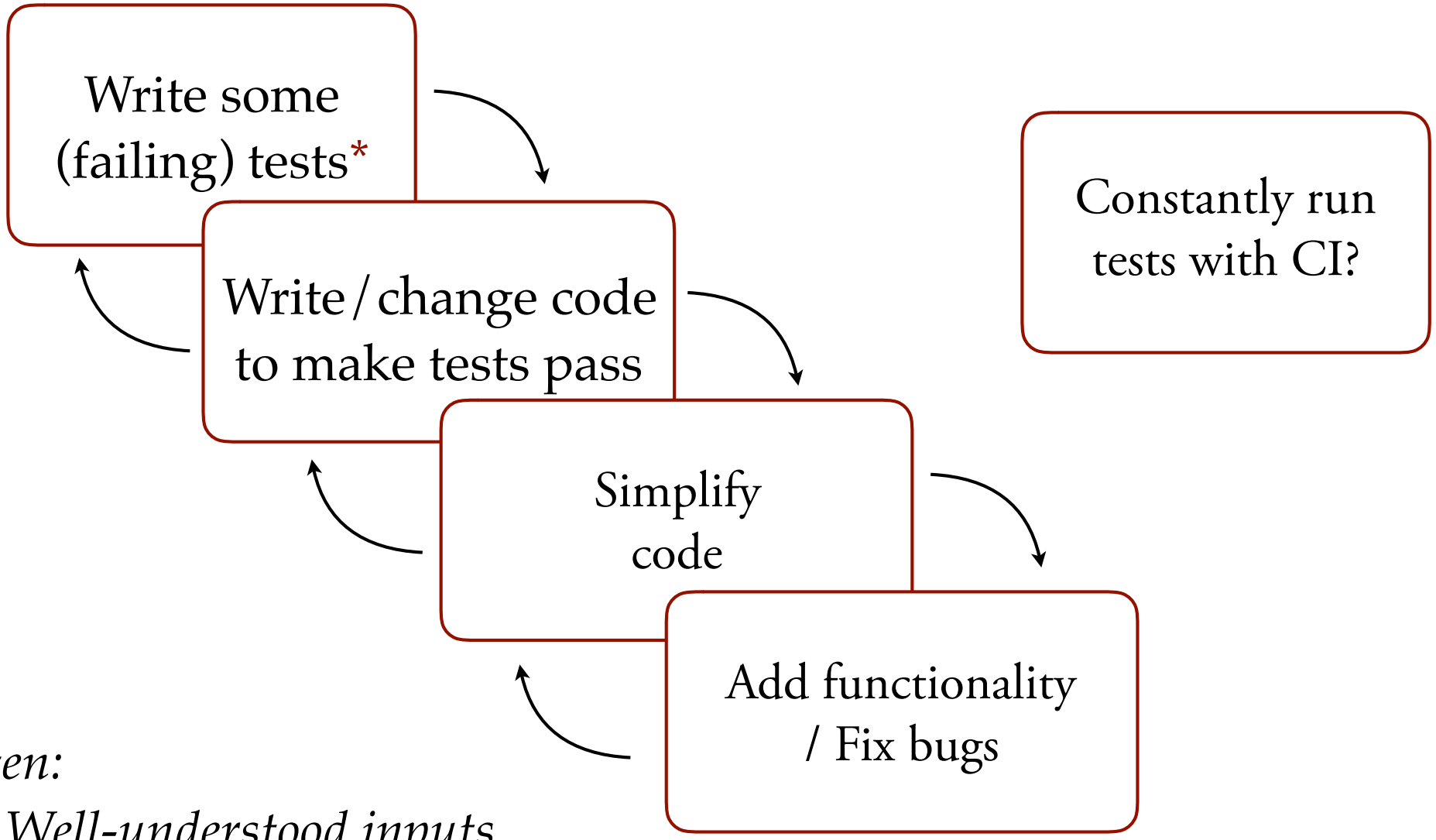
"We're not just there to help with our sponsor's confirmation bias: we're there to get it right."

— Edward Tufte, 26 September 2016

The Future of Data Analysis

https://channel9.msdn.com/Events/Machine-Learning-and-Data-Sciences-Conference/Data-Science-Summit-2016/MSDSS11

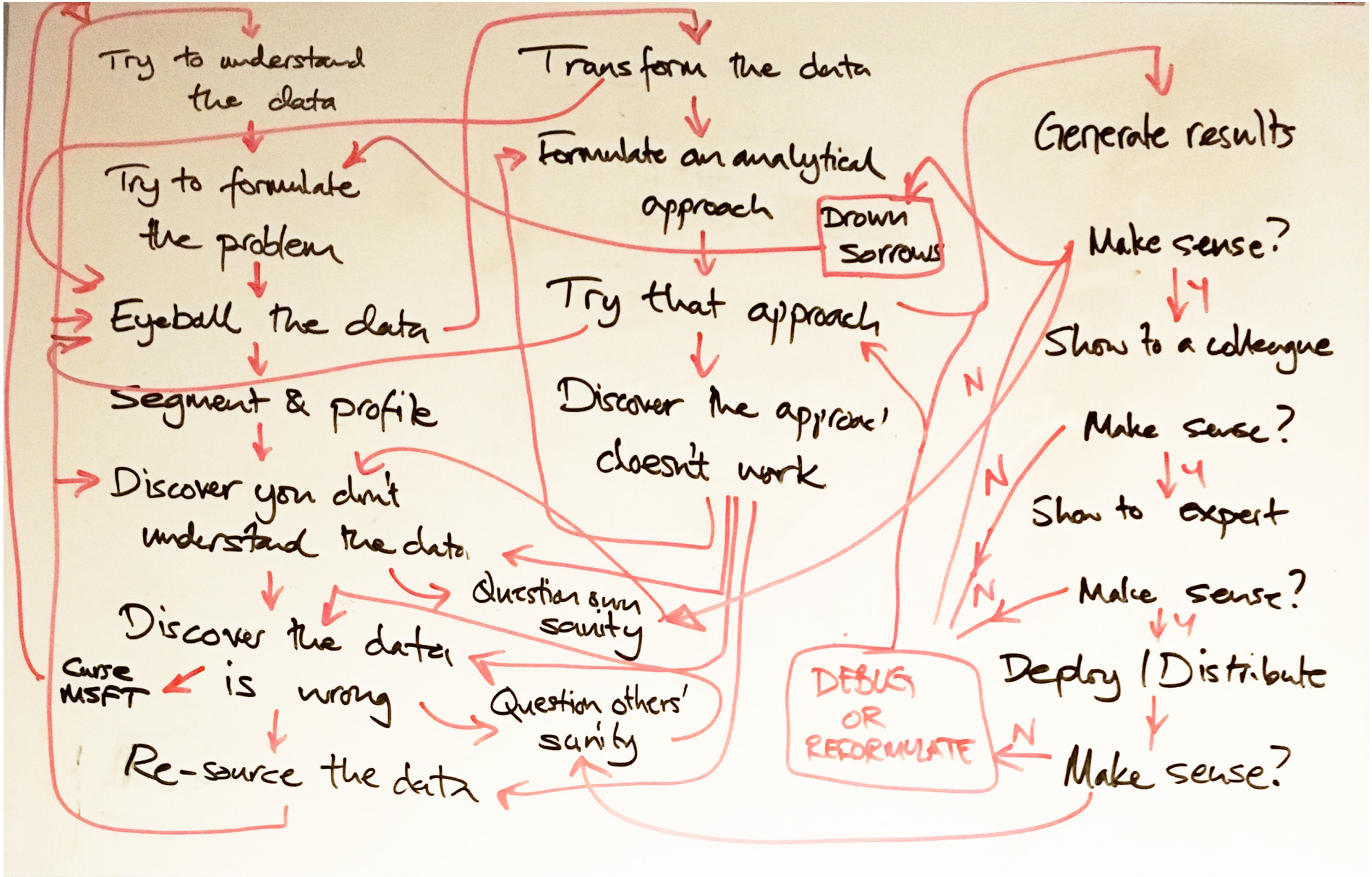
SOFTWARE DEVELOPMENT (WITH TDD)



Often:

- *Well-understood inputs*
- *Well-understood goal*
- *Many kinds of errors/failures are unmistakable*

* While mocking almost everything


Try to understand the data
Try to formulate the problem
Eyeball the data
Segment & profile
Discover you don't understand the data
Discover the data is wrong
Curse MSFT
Re-source the data
Transform the data
Formulate an analytical approach
Drown sorrows
Try that approach
Discover the approach doesn't work
Question own sanity
Question others' sanity
DEBUG OR REFORMULATE
Generate results
Make sense?
Y
Show to a colleague
Make sense?
Y
Show to expert
Make sense?
Y
Deploy / Distribute
Make sense?
N
N
N
N

Why is this
lying bastard
lying to me?

—Jeremy Paxman

"Shouldn't we do test-driven data analysis?"

— Patrick Surry, c. 2010

THE BIG IDEA

Transfer the ideas of
test-driven development
from software development
to data analysis
(mutatis mutandis)*

* *those things which need to be changed having been changed*

CORRECTNESS

STURGEON'S REVELATION

"Ninety per cent of everything is crap"

— Theodore Sturgeon, Venture, 1958

TDD ↦ TDDA

We need to extend TDD's idea of testing for

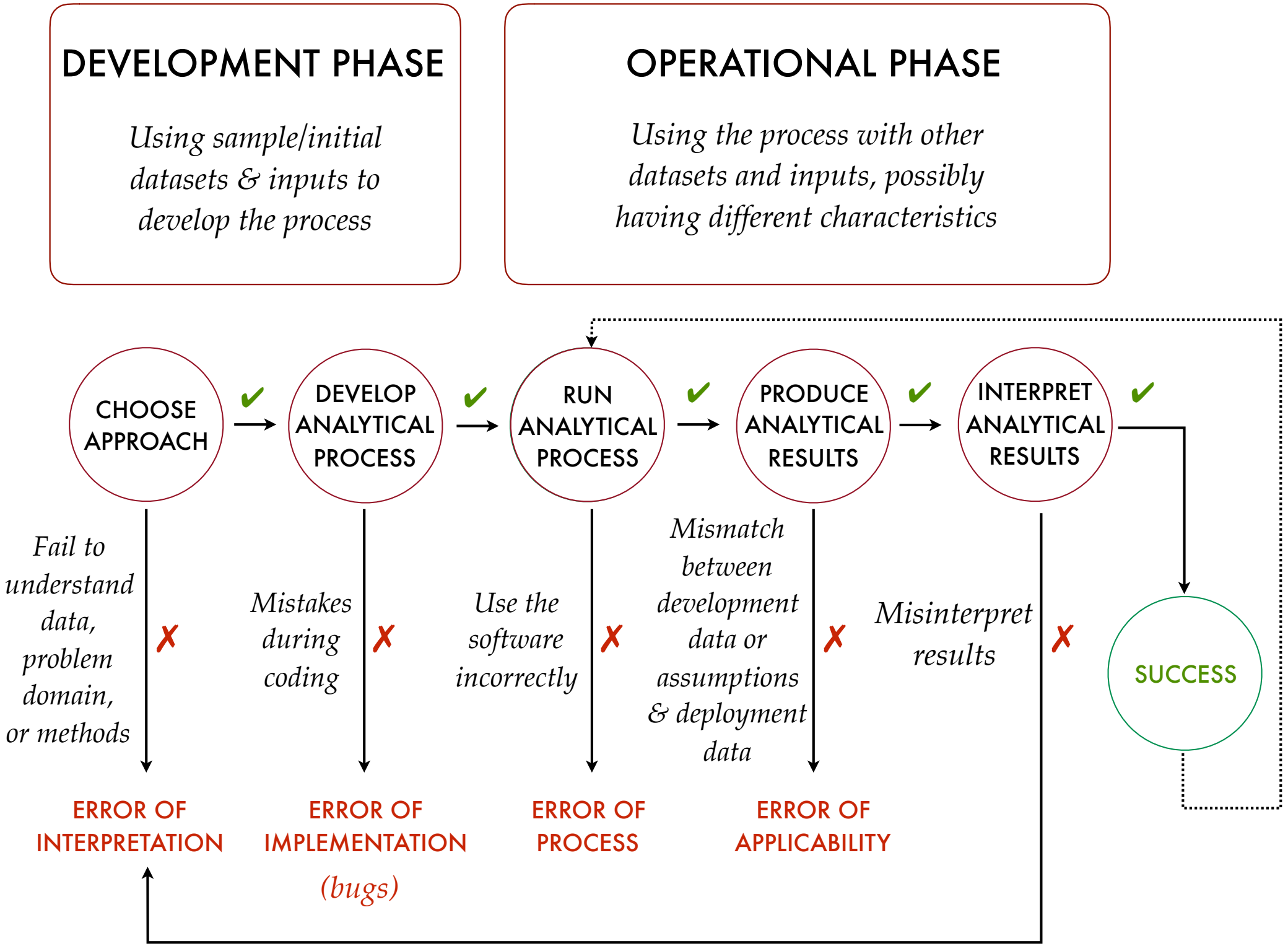
software correctness

with the idea of testing for

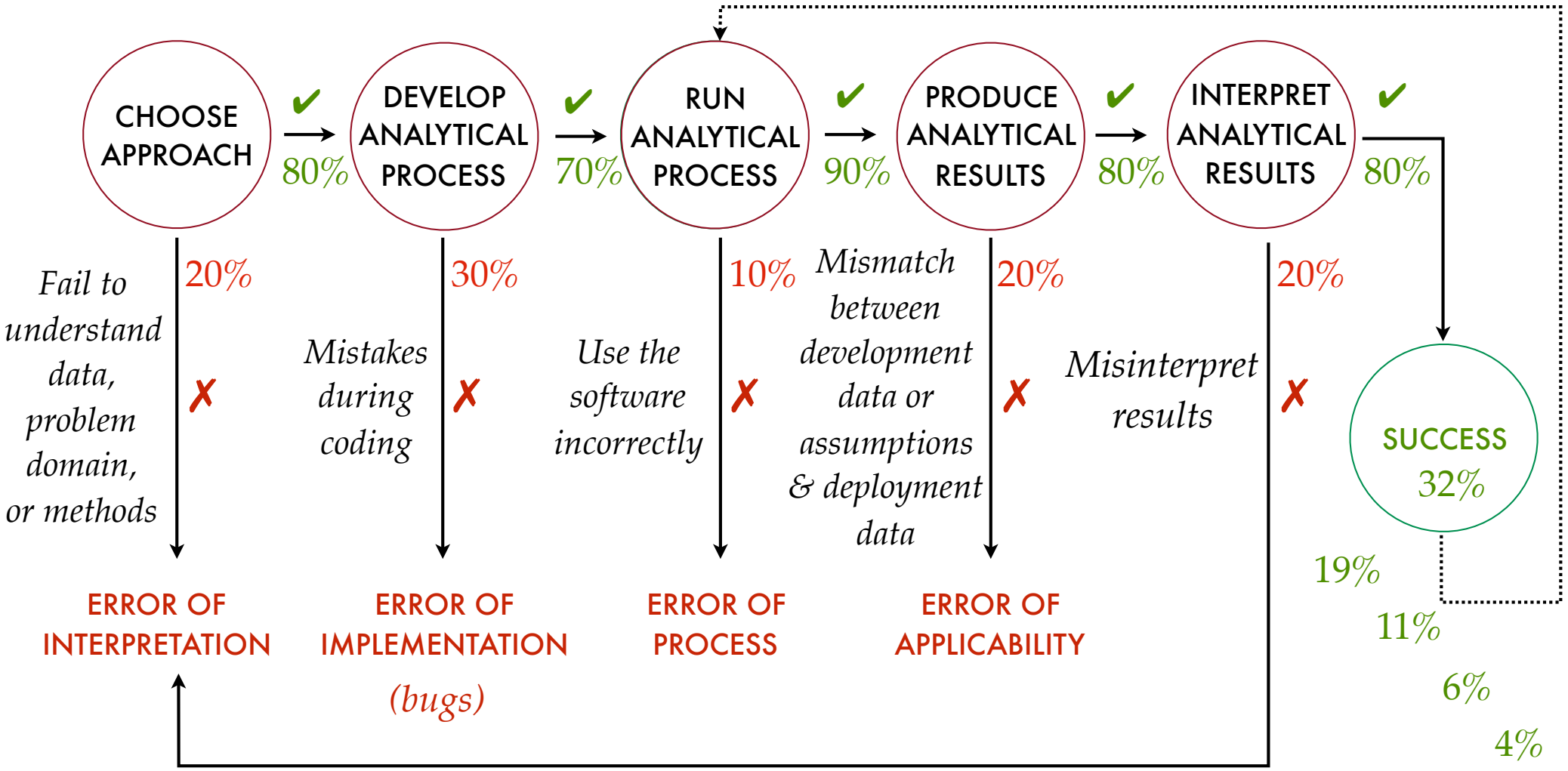
meaningfulness of analysis,

correctness and validity of input data,

& *correctness of interpretation*.


DEVELOPMENT PHASE
Using sample/initial datasets & inputs to develop the process
OPERATIONAL PHASE
Using the process with other datasets and inputs, possibly having different characteristics
CHOOSE APPROACH
DEVELOP ANALYTICAL PROCESS
RUN ANALYTICAL PROCESS
PRODUCE ANALYTICAL RESULTS
INTERPRET ANALYTICAL RESULTS
SUCCESS
Fail to understand data, problem domain, or methods
Mistakes during coding
Use the software incorrectly
Mismatch between development data or assumptions & deployment data
Misinterpret results
ERROR OF INTERPRETATION
ERROR OF IMPLEMENTATION
(bugs)
ERROR OF PROCESS
ERROR OF APPLICABILITY

If you buy into this model, it's sobering to attach probability estimates to each transition and calculate the probability of success after a few runs . . .

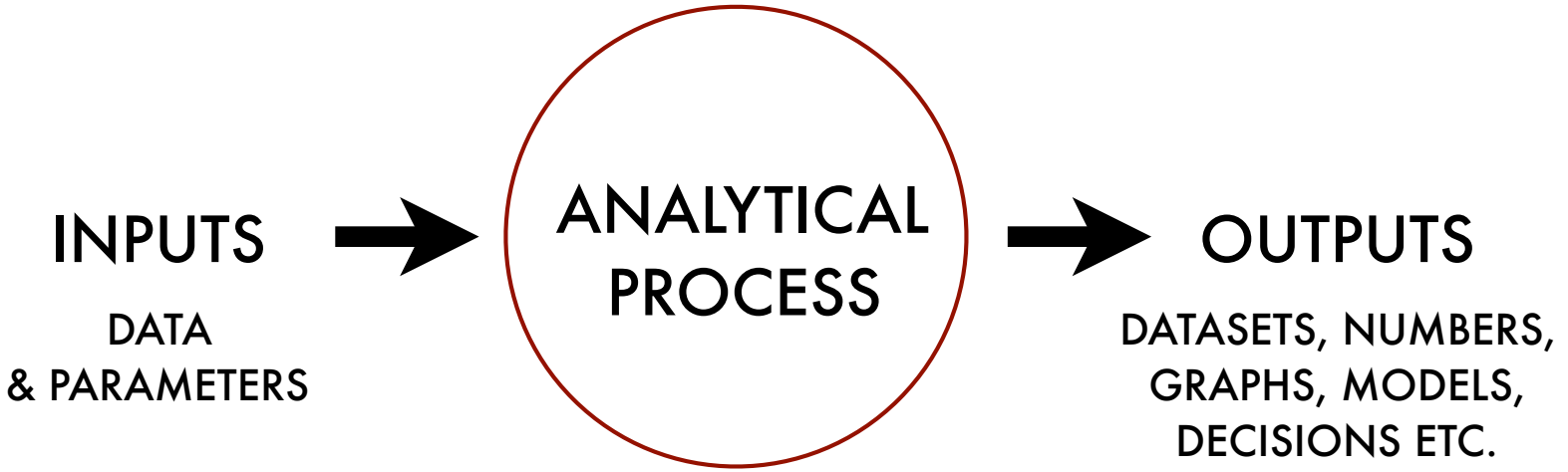


TDDA: LEVEL ZERO

REFERENCE TESTS

http://github.com/tdda/writabletestcase.py

TDDA: LEVEL ZERO



Develop a verification procedure (`diff`) *and periodically rerun:*
do the same inputs (still) produce the same outputs?

```
0 :$ python3 test_graphs.py
.....
File check failed.
Compare with "diff /Users/njr/pycon/testoutput/tsr-288-DS-5.svg /Users/njr/pycon/
reference/tsr-288-DS-5.svg".
Note exclusions: Copyright
Compare preprocessed with "diff /tmp/actual-tsr-288-DS-5.svg /tmp/expected-tsr-288-
DS-5.svg".
F
======================================================================
FAIL: testTSRGraphs (__main__.TestGraphs)
----------------------------------------------------------------------
Traceback (most recent call last):
  File "test_graphs.py", line 111, in testTSRGraphs
    maxPermutationCases=2)
  File "/Users/njr/python/tdda/writabletestcase.py", line 170, in check_file
    self.assertEqual(nFailures, 0)
AssertionError: 1 != 0

----------------------------------------------------------------------
Ran 9 tests in 11.320s

FAILED (failures=1)
```

```
1 :$ python3 test_graphs.py -W
Written /Users/njr/pycon/reference/pndate-graph.svg.
.Written /Users/njr/pycon/reference/test-pnhbar-graph.svg.
.Written /Users/njr/pycon/reference/pnd-295-SGLD-28.svg.
Written /Users/njr/pycon/reference/pnd-295-SI-1136.svg.
Written /Users/njr/pycon/reference/pnd-295-DS-235.svg.
Written /Users/njr/pycon/reference/pnd-428-BH-44.svg.
Written /Users/njr/pycon/reference/pnd-295-SI-1139.svg.
Written /Users/njr/pycon/reference/pnd-295-SGLD-8.svg.
Written /Users/njr/pycon/reference/pnd-295-BD-97.svg.
Written /Users/njr/pycon/reference/pnd-295-BD-94.svg.
Written /Users/njr/pycon/reference/pnd-428-BH-4.svg.
...
Written /Users/njr/pycon/reference/tsr-273-DS-4.svg.
Written /Users/njr/pycon/reference/tsr-278-DS-4.svg.
Written /Users/njr/pycon/reference/tsr-282-DS-3.svg.
Written /Users/njr/pycon/reference/tsr-287-DS-3.svg.
Written /Users/njr/pycon/reference/tsr-287-PW-3.svg.
Written /Users/njr/pycon/reference/tsr-288-DS-5.svg.
.
----------------------------------------------------------------------
Ran 9 tests in 11.039s

OK
```

```
0 :$ python3 test_graphs.py
.........
----------------------------------------------------------------------
Ran 9 tests in 11.658s

OK
0 :$
```

TDDA: LEVEL ONE

CONSTRAINTS

See http://github.com/tdda

/constraints/examples/pdgenerate.py

EXAMPLE CONSTRAINTS

SINGLE FIELD CONSTRAINTS	DATASET CONSTRAINTS
Age ≤ 150	The dataset must contain field **CID**
type(Age) = int	Number of records must be 118
CID ≠ NULL	One field should be tagged O
CID unique	Date should be sorted ascending
len(CardNumber) = 16	**MULTI-FIELD CONSTRAINTS**
Base in {"C", "G", "A", T"}	StartDate ≤ EndDate
Vote ≠ "Trump"	AlmostEqual(F, m * a, 6)
StartDate < tomorrow()	sum(Favourite*) = 1
v < 2.97e10	minVal ≤ medianVal ≤ maxVal
Height ~ N(1.8, 0.2)	V ≤ H * w * d

Z	Name	Symbol	Period	Group	Chemical Series	Atomic Weight	Etymology	Relative Atomic Mass	Melting Point C	Melting Point K	Boiling Point C	Boiling Point F	Density	Description	Colour
1	Hydrogen	H	1	1	Nonmetal	1.01	Greek hydrogenes	1.01	-258.98	14.20	-252.87	-423.17	0.00	gas	colorless
2	Helium	He	1	18	Noble gas	4.00	Greek helios	4.00	⌀	⌀	-268.93	-452.07	0.00	noble gas	⌀
3	Lithium	Li	2	1	Alkali metal	6.94	Greek lithos	6.94	180.70	453.90	1,342.00	2,448.00	0.53	0.6	silvery white/gray
4	Beryllium	Be	2	2	Alkaline earth metal	9.01	beryl	9.01	1,287.00	1,560.00	2,469.00	4,476.00	1.85	5.5	⌀
5	Boron	B	2	13	Metalloid	10.81	borax	10.81	2,300.00	2,570.00	3,927.00	7,101.00	2.34	9.3	black/brown/ amorphous boron is a brown powder, metallic boron is black
6	Carbon	C	2	14	Nonmetal	12.01	Latin carbo	12.01	3,675.00	3,948.00	4,027.00	7,280.00	2.27	1-2 graphite	black
7	Nitrogen	N	2	15	Nonmetal	14.01	Greek nitron	14.01	-209.86	63.30	-195.79	-320.42	0.00	gas	⌀
8	Oxygen	O	2	16	Nonmetal	16.00	Greek oxys	16.00	-222.65	50.50	-182.95	-297.31	0.00	gas	⌀
9	Fluorine	F	2	17	Halogen	19.00	Latin fluo	19.00	-219.52	53.60	-188.12	-306.62	0.00	halogen gas	yellow-green or yellowish brown
10	Neon	Ne	2	18	Noble gas	20.18	Greek neos	20.18	-248.45	24.70	-246.08	-410.94	0.00	noble gas	⌀
11	Sodium	Na	3	1	Alkali metal	22.99	Latin natrium	22.99	98.00	371.00	883.00	1,621.00	0.97	0.5	waxy, silvery white
12	Magnesium	Mg	3	2	Alkaline earth metal	24.31	Magnesia, Greece	24.31	650.00	923.00	1,090.00	1,994.00	1.74	2.5	silvery metallic
13	Aluminium	Al	3	13	Poor metal	26.98	Latin alumen	26.98	660.25	933.40	2,519.00	4,566.00	2.70	2.75	silvery
14	Silicon	Si	3	14	Metalloid	28.09	Latin silex	28.09	1,410.00	1,680.00	3,265.00	5,909.00	2.33	6.5 metalloid	dark gray, bluish tinge
15	Phosphorus	P	3	15	Nonmetal	30.97	Greek phosphoros	30.97	44.10	317.30	280.00	536.00	1.82	nonmetal	waxy white/ red/ black/ colorless
16	Sulfur	S	3	16	Nonmetal	32.07	Latin sulfur	32.07	115.36	388.51	444.60	832.30	2.07	2	lemon yellow
17	Chlorine	Cl	3	17	Halogen	35.45	Greek chloros	35.45	-100.84	172.00	-34.04	-29.27	0.00	halogen gas	yellowish green or greenish yellow
18	Argon	Ar	3	18	Noble gas	39.95	Greek argon	39.95	-189.19	84.00	-185.85	-302.53	0.00	noble gas	⌀
19	Potassium	K	4	1	Alkali metal	39.10	Latin kalium	39.10	63.35	336.50	759.00	1,398.00	0.86	0.4	silvery white
20	Calcium	Ca	4	2	Alkaline earth metal	40.08	Latin calx	40.08	839.00	1,112.00	1,484.00	2,703.00	1.54	1.75	gray

Miró

Data analysis (lisp-like) interface

This is Miró version 2.7.80, running under Dalí 1.0.31.


```
Seed: 1520109300
Logs started at 2016/10/31 16:00:40 host godel.local.
Logging to /Users/njr/miro/log/2016/10/31/session005.

elements.feather     elements.miro        elements92.feather   elements92.miro

[1]>  load elements92

elements92.feather: 92 records; 92 (100%) selected; 16 fields.

[2]>  discover -t -o /tmp/elements92
```

Individual Field Constraints							
Name	Type Allowed	Min Allowed	Max Allowed	Sign Allowed	Nulls Allowed	Duplicates Allowed	Values Allowed
Z	int	1	92	positive	0	no	
Name	string	length 3	length 12		0	no	
Symbol	string	length 1	length 2		0	no	
Period	int	1	7	positive	0	-	
Group	int	1	18	positive		-	
ChemicalSeries	string	length 7	length 20		0	-	"Actinoid" "Alkali metal" "Alkaline earth metal" "Halogen" "Lanthanoid" "Metalloid" "Noble gas" "Nonmetal" "Poor metal" "Transition metal"
AtomicWeight	real	1.007946	238.028913	positive	0		
Etymology	string	length 4	length 39		0	-	
RelativeAtomicMass	real	1.007946	238.028913	positive	0		
MeltingPointC	real	-258.975000	3675.0		1		
MeltingPointKelvin	real	14.200000	3948.0	positive	1		
BoilingPointC	real	-268.930000	5596.0		0		
BoilingPointF	real	-452.070000	10105.0		0		
Density	real	0.000089	22.610000	positive	0		
Description	string	length 1	length 83			-	
Colour	string	length 4	length 80			-	

/tmp/elements92.tdda


```
{
   "fields": {
      "Z": {"type": "int", "min": 1, "max": 92, "sign": "positive", "max_nulls": 0, "no_duplicates": true},
      "Name": {"type": "string", "min_length": 3, "max_length": 12, "max_nulls": 0, "no_duplicates": true},
      "Symbol": {"type": "string", "min_length": 1,"max_length": 2, "max_nulls": 0, "no_duplicates": true},
      "Period": {"type": "int", "min": 1, "max": 7, "sign": "positive", "max_nulls": 0},
      "Group": {"type": "int", "min": 1, "max": 18,"sign": "positive"},
      "ChemicalSeries": {"type": "string", "min_length": 7, "max_length": 20, "max_nulls": 0,
                         "allowed_values": ["Actinoid", "Alkali metal", "Alkaline earth metal",
                                            "Halogen", "Lanthanoid", "Metalloid", "Noble gas",
                                            "Nonmetal", "Poor metal", "Transition metal"]},
      "AtomicWeight": {"type": "real", "min": 1.007946, "max": 238.028914, "sign": "positive", "max_nulls": 0},
      "Etymology": {"type": "string", "min_length": 4, "max_length": 39, "max_nulls": 0},
      "RelativeAtomicMass": {"type": "real", "min": 1.007946, "max": 238.028914, "sign": "positive",
                             "max_nulls": 0},
      "MeltingPointC": {"type": "real", "min": -258.975, "max": 3675.0, "max_nulls": 1},
      "MeltingPointKelvin": {"type": "real", "min": 14.2, "max": 3948.0, "sign": "positive", "max_nulls": 1},
      "BoilingPointC": {"type": "real", "min": -268.93, "max": 5596.0, "max_nulls": 0},
      "BoilingPointF": {"type": "real", "min": -452.07, "max": 10105.0, "max_nulls": 0},
      "Density": {"type": "real", "min": 8.9e-05, "max": 22.610001, "sign": "positive", "max_nulls": 0},
      "Description": {"type": "string", "min_length": 1, "max_length": 83},
      "Colour": {"type": "string", "min_length": 4, "max_length": 80}
   }
}
```

Miró

Data analysis (lisp-like) interface

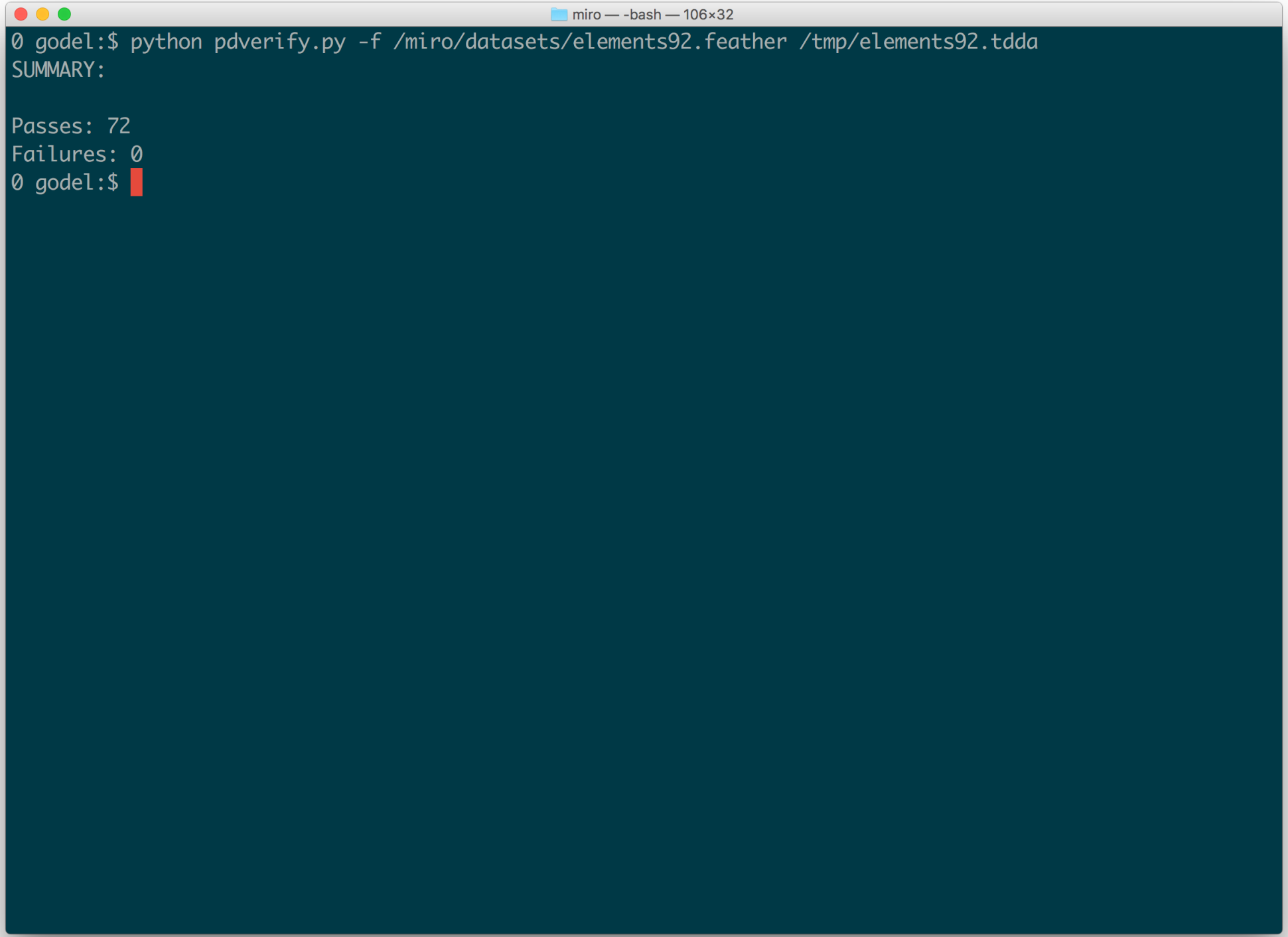
This is Miró version 2.7.80, running under Dalí 1.0.31.

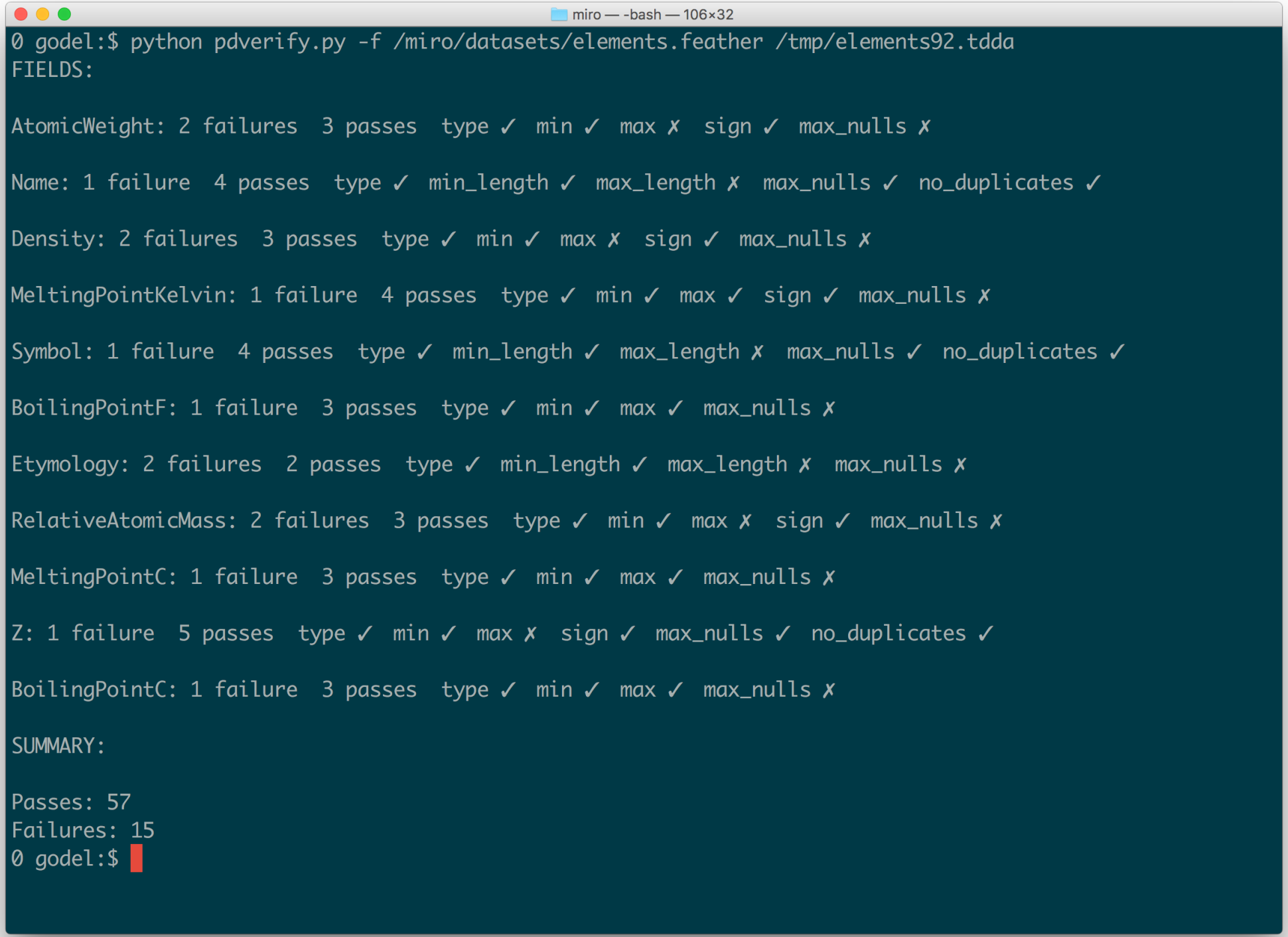

```
Seed: 1507023627
Logs started at 2016/10/31 16:27:05 host godel.local.
Logging to /Users/njr/miro/log/2016/10/31/session015.

[1]>  load elements92

elements92.feather: 92 records; 92 (100%) selected; 16 fields.

[2]>  verify -f /tmp/elements92.tdda

No failures

[3]>  load elements

elements.feather: 118 records; 118 (100%) selected; 16 fields.

[4]>  verify -f /tmp/elements92.tdda
```

Miro Warning: 15 failures during constraint verification for dataset elements.feather

Name	Failures	Type			Minimum			Maximum			Sign			Max Nulls			Duplicates				Values		
		Allowed	Actual	✓	Allowed	Actual	✓	Allowed	Actual	✓	Allowed	Actual	✓	Allowed	Actual	✓	Allowed	Actual	Non-Null Values	✓	Allowed	Actual	✓
AtomicWeight	2	real	real	✓	1.01	1.01	✓	238.03	294.00	✗	positive	positive	✓	0	1	✗	-		-	-	-	-	-
Etymology	2	string	string	✓	length 4	length 4	✓	length 39	length 53	✗	-	-	-	0	1	✗	-	3	117	-	-	-	-
RelativeAtomicMass	2	real	real	✓	1.01	1.01	✓	238.03	294.00	✗	positive	positive	✓	0	1	✗	-		-	-	-	-	-
Density	2	real	real	✓	0.00	0.00	✓	22.61	41.00	✗	positive	positive	✓	0	5	✗	-		-	-	-	-	-
Z	1	int	int	✓	1	1	✓	92	118	✗	positive	positive	✓	0	0	✓	no	0	118	✓	-	-	-
Name	1	string	string	✓	length 3	length 3	✓	length 12	length 13	✗	-	-	-	0	0	✓	no	0	118	✓	-	-	-
Symbol	1	string	string	✓	length 1	length 1	✓	length 2	length 3	✗	-	-	-	0	0	✓	no	0	118	✓	-	-	-
MeltingPointC	1	real	real	✓	-258.98	-258.98	✓	3,675.00	3,675.00	✓	-	-	-	1	20	✗	-		-	-	-	-	-
MeltingPointKelvin	1	real	real	✓	14.20	14.20	✓	3,948.00	3,948.00	✓	positive	positive	✓	1	20	✗	-		-	-	-	-	-
BoilingPointC	1	real	real	✓	-268.93	-268.93	✓	5,596.00	5,596.00	✓	-	-	-	0	20	✗	-		-	-	-	-	-
BoilingPointF	1	real	real	✓	-452.07	-452.07	✓	10,105.00	10,105.00	✓	-	-	-	0	20	✗	-		-	-	-	-	-

15 failures


miro — -bash — 106×32
0 godel:$ python pdverify.py -f /miro/datasets/elements92.feather /tmp/elements92.tdda
SUMMARY:
Passes: 72
Failures: 0
0 godel:$


miro — -bash — 106×32
0 godel:$ python pdverify.py -f /miro/datasets/elements.feather /tmp/elements92.tdda
FIELDS:
AtomicWeight: 2 failures 3 passes type ✓ min ✓ max ✗ sign ✓ max_nulls ✗
Name: 1 failure 4 passes type ✓ min_length ✓ max_length ✗ max_nulls ✓ no_duplicates ✓
Density: 2 failures 3 passes type ✓ min ✓ max ✗ sign ✓ max_nulls ✗
MeltingPointKelvin: 1 failure 4 passes type ✓ min ✓ max ✓ sign ✓ max_nulls ✗
Symbol: 1 failure 4 passes type ✓ min_length ✓ max_length ✗ max_nulls ✓ no_duplicates ✓
BoilingPointF: 1 failure 3 passes type ✓ min ✓ max ✓ max_nulls ✗
Etymology: 2 failures 2 passes type ✓ min_length ✓ max_length ✗ max_nulls ✗
RelativeAtomicMass: 2 failures 3 passes type ✓ min ✓ max ✗ sign ✓ max_nulls ✗
MeltingPointC: 1 failure 3 passes type ✓ min ✓ max ✓ max_nulls ✗
Z: 1 failure 5 passes type ✓ min ✓ max ✗ sign ✓ max_nulls ✓ no_duplicates ✓
BoilingPointC: 1 failure 3 passes type ✓ min ✓ max ✓ max_nulls ✗
SUMMARY:
Passes: 57
Failures: 15
0 godel:$

pdverifyelements1.py

```
# pdverifyelements1.py

from __future__ import print_function
from tdda.constraints.pdconstraints import verify_df
from pmmif.featherpmm import read_dataframe

ds = read_dataframe('/miro/datasets/elements92.feather')
v = verify_df(ds.df, '/tmp/elements92.tdda')
print('Passes:', v.passes, 'Failures:', v.failures)
```

OUTPUT

```
$ python pdverifyelements1.py
Passes: 72 Failures: 0
```

pdverifyelements2.py

```
# pdverifyelements2.py

from __future__ import print_function
from tdda.constraints.pdconstraints import verify_df
from pmmif.featherpmm import read_dataframe

ds = read_dataframe('/miro/datasets/elements.feather')
v = verify_df(ds.df, '/tmp/elements92.tdda', report='fields')
print(str(v))

dfv = v.to_frame()
print(dfv[dfv.failures > 0])
```

OUTPUT 1

```
$ python pdverifyelements2.py
FIELDS:

AtomicWeight: 2 failures  3 passes  type ✓  min ✓  max ✗  sign ✓  max_nulls ✗

Name: 1 failure  4 passes  type ✓  min_length ✓  max_length ✗  max_nulls ✓  no_duplicates ✓

Density: 2 failures  3 passes  type ✓  min ✓  max ✗  sign ✓  max_nulls ✗

MeltingPointKelvin: 1 failure  4 passes  type ✓  min ✓  max ✓  sign ✓  max_nulls ✗

Symbol: 1 failure  4 passes  type ✓  min_length ✓  max_length ✗  max_nulls ✓  no_duplicates ✓

BoilingPointF: 1 failure  3 passes  type ✓  min ✓  max ✓  max_nulls ✗

Etymology: 2 failures  2 passes  type ✓  min_length ✓  max_length ✗  max_nulls ✗

RelativeAtomicMass: 2 failures  3 passes  type ✓  min ✓  max ✗  sign ✓  max_nulls ✗

MeltingPointC: 1 failure  3 passes  type ✓  min ✓  max ✓  max_nulls ✗

Z: 1 failure  5 passes  type ✓  min ✓  max ✗  sign ✓  max_nulls ✓  no_duplicates ✓

BoilingPointC: 1 failure  3 passes  type ✓  min ✓  max ✓  max_nulls ✗

SUMMARY:

Passes: 57
Failures: 15
```

OUTPUT 2 (DataFrame)

	field	failures	passes	type	min	min_length	max	\
0	AtomicWeight	2	3	True	True	NaN	False	
2	Name	1	4	True	NaN	True	NaN	
3	Density	2	3	True	True	NaN	False	
4	MeltingPointKelvin	1	4	True	True	NaN	True	
5	Symbol	1	4	True	NaN	True	NaN	
7	BoilingPointF	1	3	True	True	NaN	True	
8	Etymology	2	2	True	NaN	True	NaN	
9	RelativeAtomicMass	2	3	True	True	NaN	False	
11	MeltingPointC	1	3	True	True	NaN	True	
12	Z	1	5	True	True	NaN	False	
13	BoilingPointC	1	3	True	True	NaN	True	

	max_length	sign	max_nulls	no_duplicates	allowed_values
0	NaN	True	False	NaN	NaN
2	False	NaN	True	True	NaN
3	NaN	True	False	NaN	NaN
4	NaN	True	False	NaN	NaN
5	False	NaN	True	True	NaN
7	NaN	NaN	False	NaN	NaN
8	False	NaN	False	NaN	NaN
9	NaN	True	False	NaN	NaN
11	NaN	NaN	False	NaN	NaN
12	NaN	True	True	True	NaN
13	NaN	NaN	False	NaN	NaN

OUTPUT of to_frame()

	field	failures	passes	type	min	min_length	max	\
0	AtomicWeight	2	3	True	True	NaN	False	
2	Name	1	4	True	NaN	True	NaN	
3	Density	2	3	True	True	NaN	False	
4	MeltingPointKelvin	1	4	True	True	NaN	True	
5	Symbol	1	4	True	NaN	True	NaN	
7	BoilingPointF	1	3	True	True	NaN	True	
8	Etymology	2	2	True	NaN	True	NaN	
9	RelativeAtomicMass	2	3	True	True	NaN	False	
11	MeltingPointC	1	3	True	True	NaN	True	
12	Z	1	5	True	True	NaN	False	
13	BoilingPointC	1	3	True	True	NaN	True	

	max_length	sign	max_nulls	no_duplicates	allowed_values
0	NaN	True	False	NaN	NaN
2	False	NaN	True	True	NaN
3	NaN	True	False	NaN	NaN
4	NaN	True	False	NaN	NaN
5	False	NaN	True	True	NaN
7	NaN	NaN	False	NaN	NaN
8	False	NaN	False	NaN	NaN
9	NaN	True	False	NaN	NaN
11	NaN	NaN	False	NaN	NaN
12	NaN	True	True	True	NaN
13	NaN	NaN	False	NaN	NaN

CONSTRAINTS

True *Satisfied*

FALSE *Not satisfied*

NaN *No constraint*

pdgenerate.py

```
# pdgenerate.py

from __future__ import print_function
import pandas as pd

from tdda.constraints.pdconstraints import discover_constraints

df = pd.read_csv('../testdata/elements92.csv')
constraints = discover_constraints(df)
with open('elements92.tdda', 'w') as f:
    f.write(constraints.to_json())
print('Written elements92.tdda')

print(verify_df(df, 'elements92.tdda'))
```

OUTPUT of pdgenerate.py

```
{
    "fields": {
        "AtomicWeight": {
            "type": "real",
            "min": 1.007946999999999,
            "max": 238.02891299999999,
            "sign": "positive",
            "max_nulls": 0
        },
        "Group": {
            "type": “real",
.
.
.
SUMMARY:

Passes: 72
Failures: 0
```

ABSENT CONSTRAINTS

Gregory (Scotland Yard detective): *"Is there any other point to which you would wish to draw my attention?"*
Holmes: *"To the curious incident of the dog in the night-time."*
Gregory: *"The dog did nothing in the night-time."*
Holmes: *"That was the curious incident."*

— *Silver Blaze*, in *Memoirs of Sherlock Holmes*
Arthur Conan Doyle, 1892.

SUMMARY

LEVEL ZERO *Reference tests (writabletestcase.py)*

- *Support for sophisticated checking*
- *Useful differences*
- *Rewriting output verified as (now) correct*

LEVEL ONE *Input/Intermediate/Output Dataset Constraints*

- *Automatic constraint discovery*
- *Systematic constraint verification*

njr@StochasticSolutions.com

http://tdda.info

http://github.com/tdda

@tdda0 @njr0

Correct interpretation: Zero

Error of interpretation: Letter "Oh"

OK, everything you need to know about TEX has been explained—unless you happen to be fallible. If you don't plan to make any errors, don't bother to read this chapter.

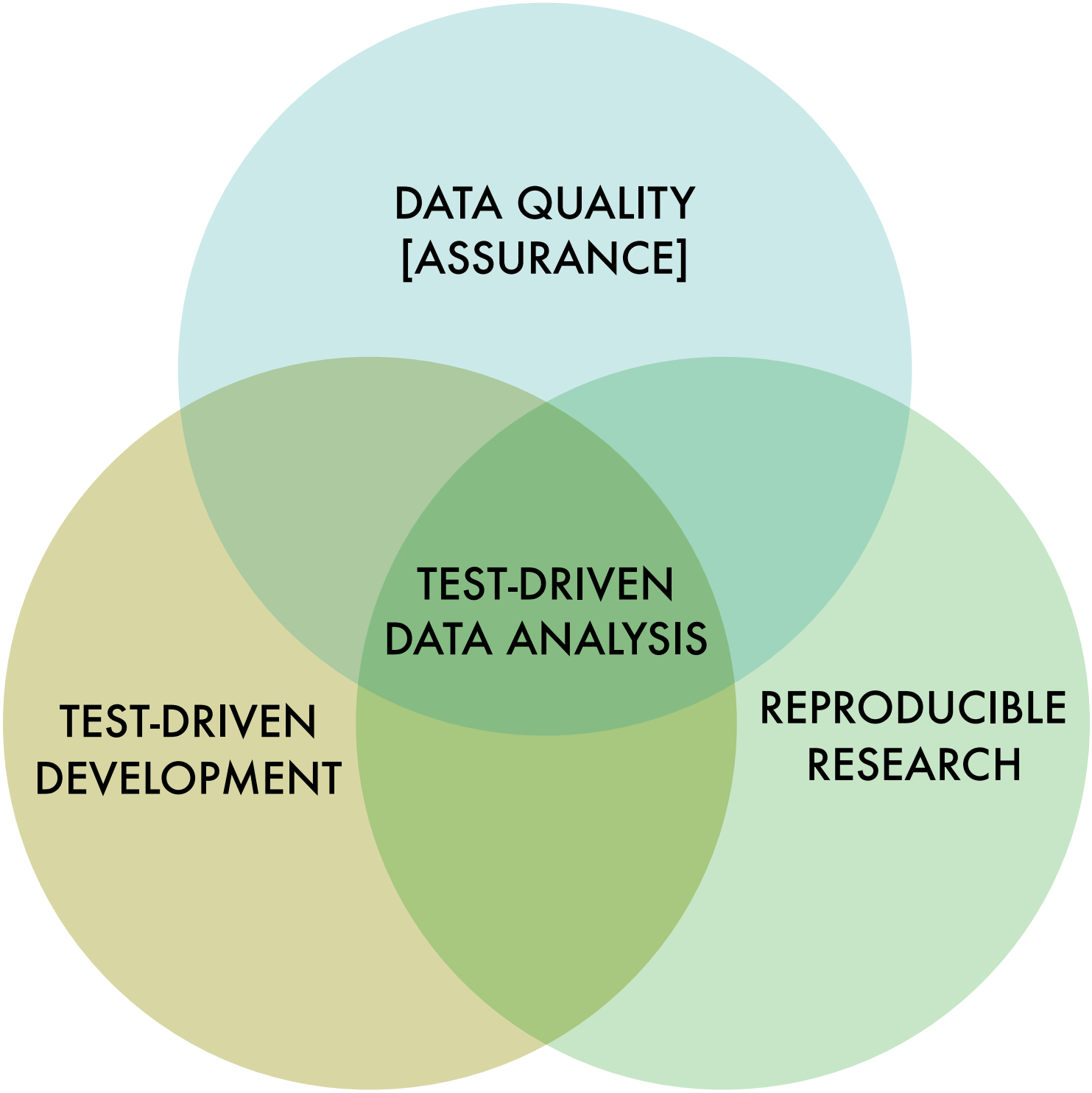
— Donald E. Knuth, author TEX

Chapter 27, The TEXbook,

Recovery from Errors

ANALYTICAL SNAFU

Situation normal
— analysis fouled up.*


DATA QUALITY
[ASSURANCE]
TEST-DRIVEN
DATA ANALYSIS
TEST-DRIVEN
DEVELOPMENT
REPRODUCIBLE
RESEARCH