

# Contracts Data Mining (CDM)

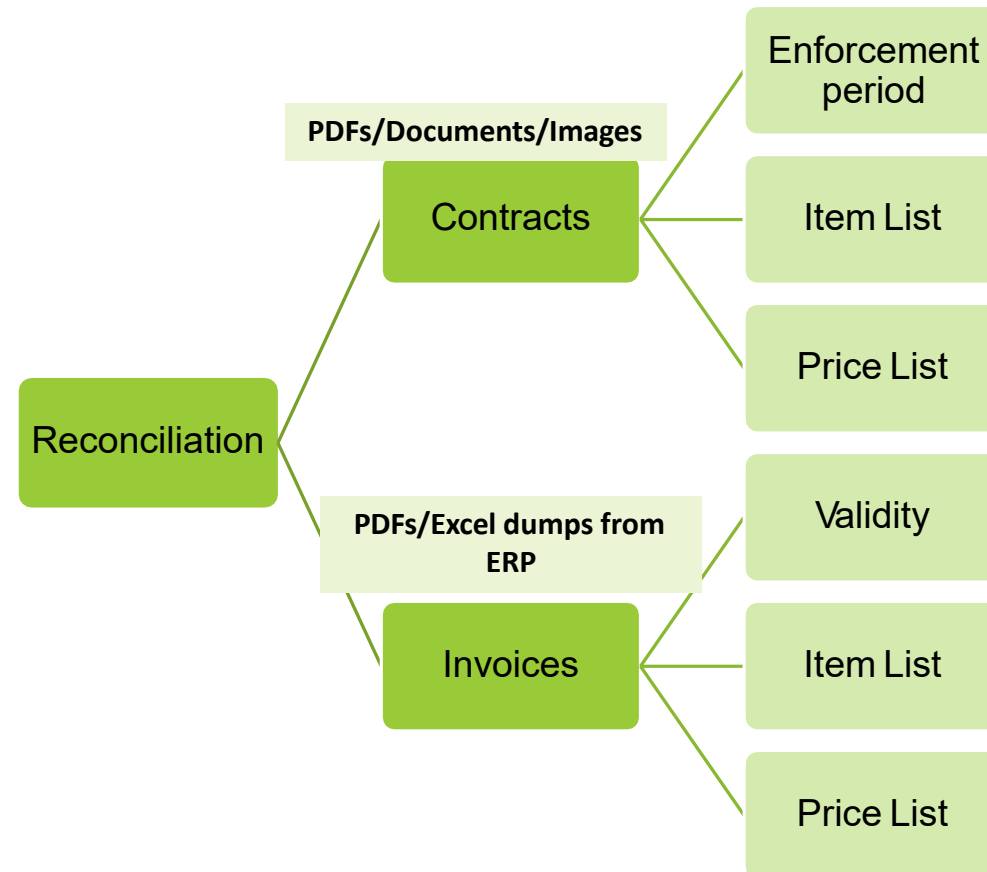
## “Audits of the Future”

---

Reduce your revenue leakage using Data Mining of Contract Documents with Invoices

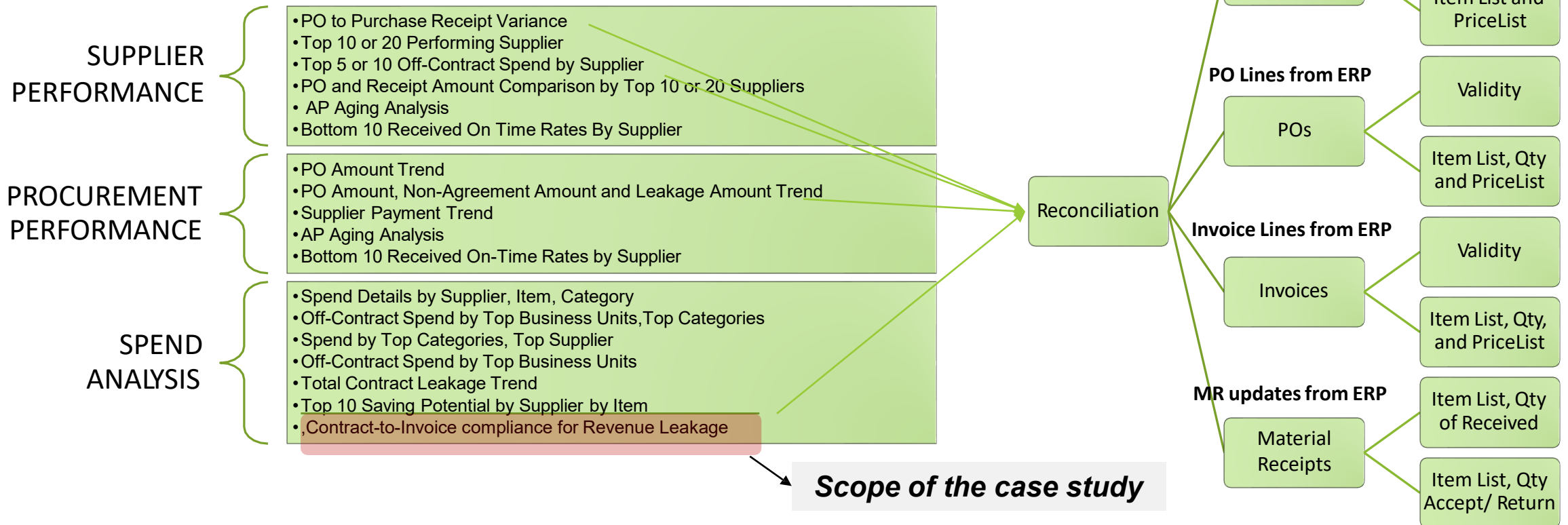
# Need for data mining of Contract Documents

- Contract documents are typically in form of PDF/word templates and stored either in a contract management system (CMS) /a document management system (DMS)
- Inconsistent terms and conditions of contracts
- Lack of standardized contract templates
- Possibility of booking error rate (15% – 25%)

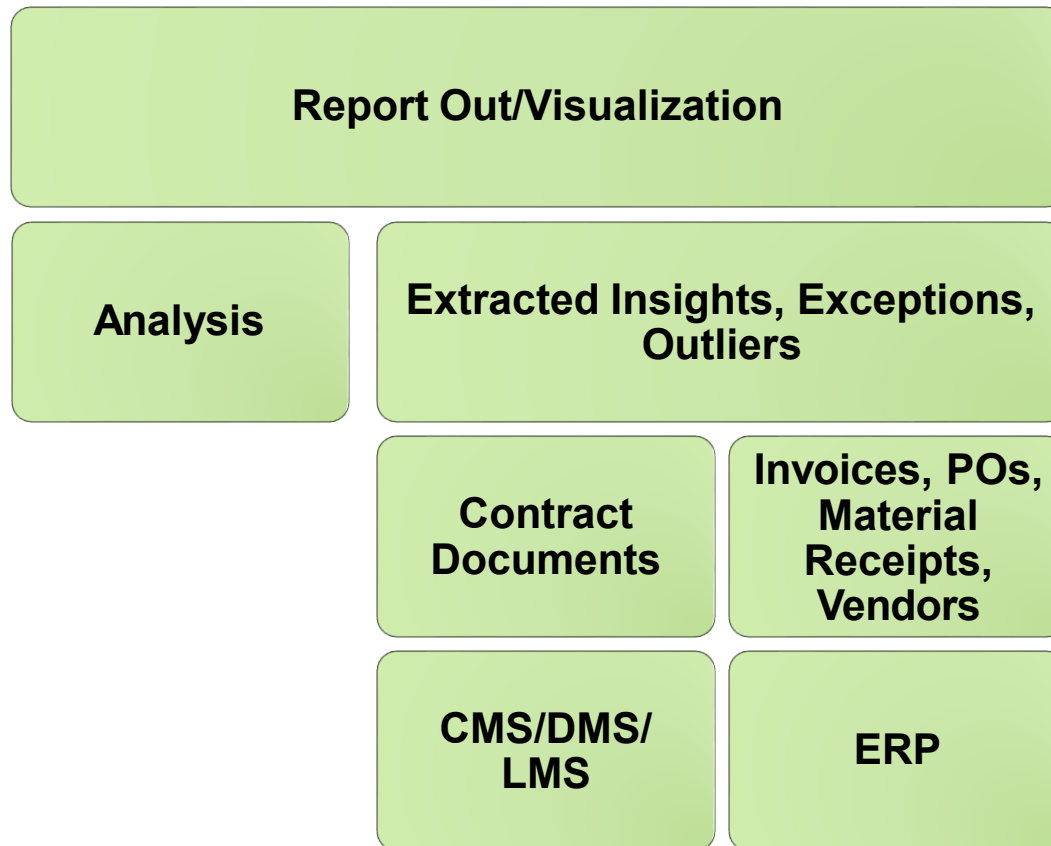


# Moving towards a 4-way matching

## Performance Data Discovery

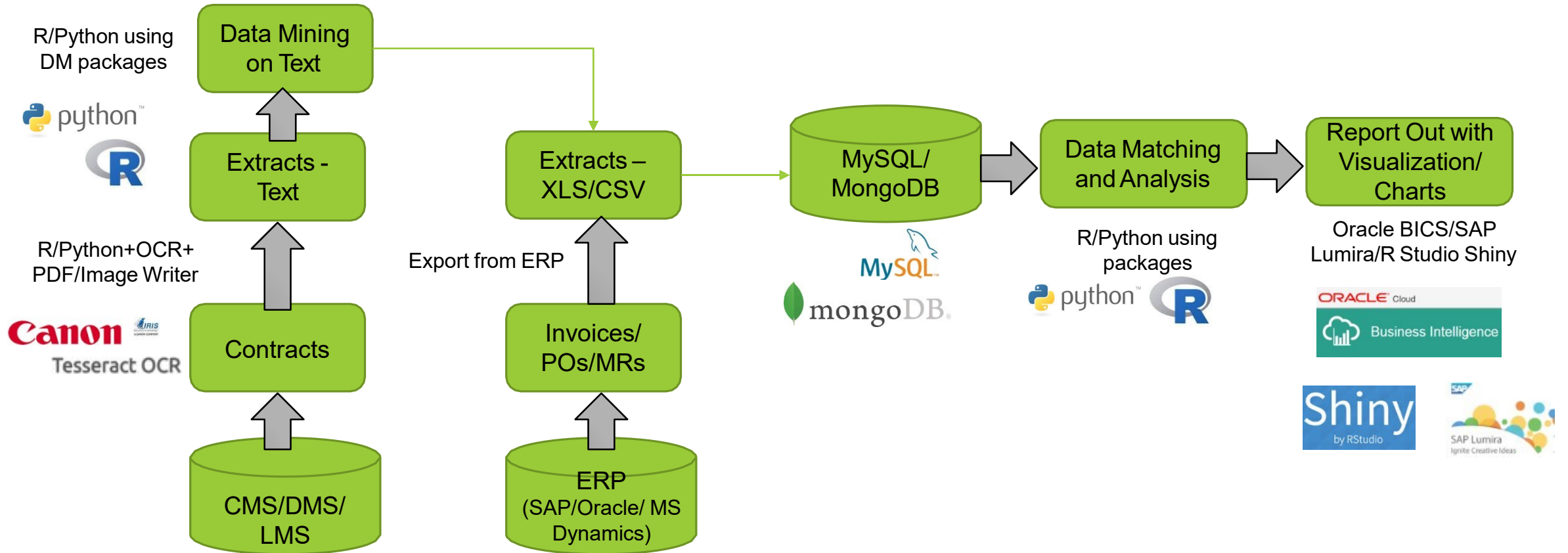


# APA - Functional Architecture



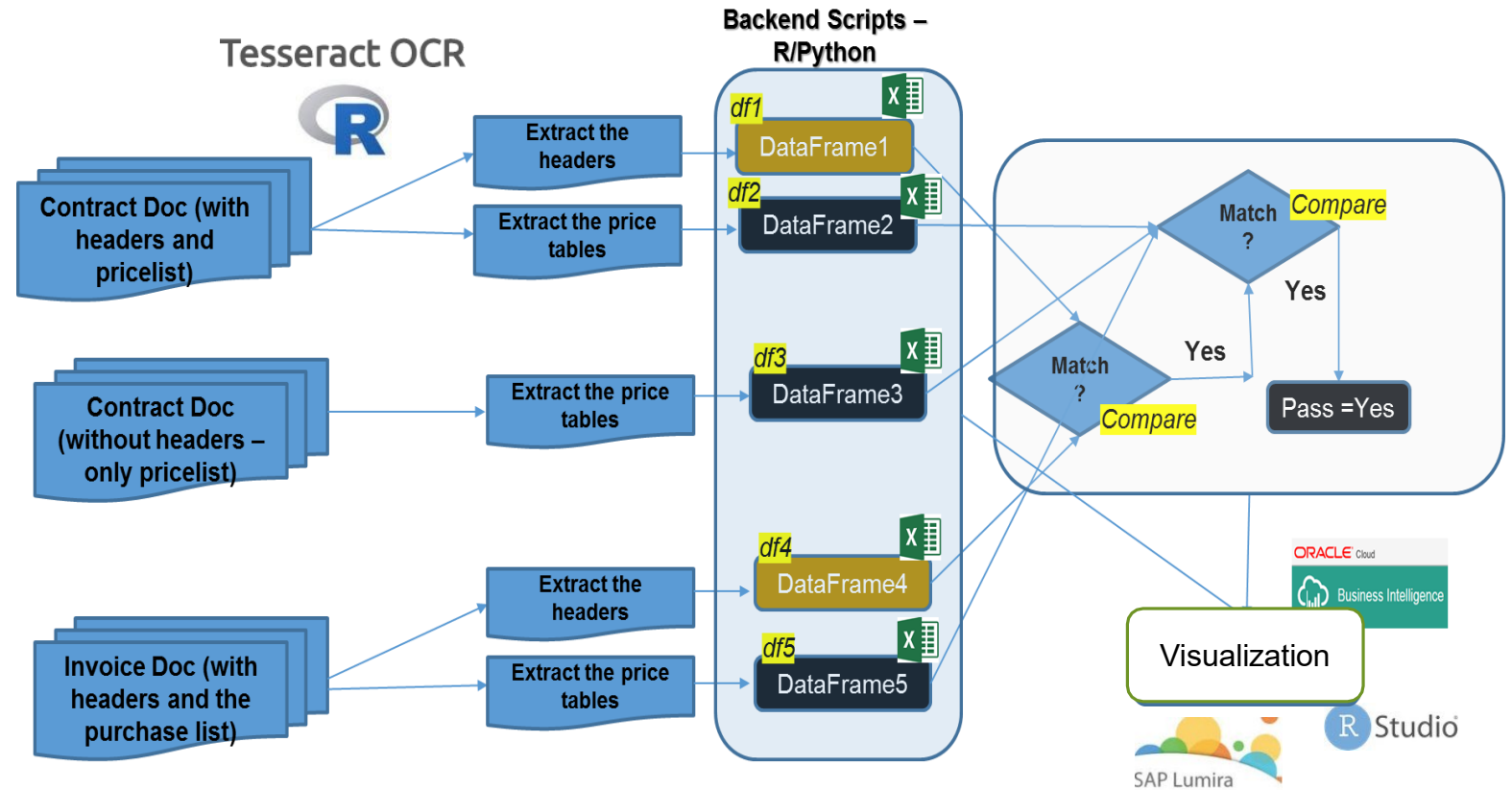
- 1) **ERP** – SAP/Oracle/MS Dynamics
- 2) **CMS/DMS/LMS** – 3rd party or part of ERP system
- 3) **Contact Documents** – Images/PDFs/Word Docs
- 4) **OCR** – Open Source (Google), Commercial (Customer Preference/Canon)
- 5) **Invoices/POs/MRs** – ERP dumps in xls/csv formats
- 6) **Data Mining on Extracted Insights** – R/Python
- 7) **Report out/Visualization** – Oracle BICS/SAP Lumira/R Shiny

# APA - Technical Architecture



# Data Extraction and Matching from Contracts and Invoices

- Automate your audit process by reconciling the contracts and invoices
- Data Quality Management using open source (Google)/commercial OCR
- Interactive visualizations from your loaded unstructured forms of contracts and invoices



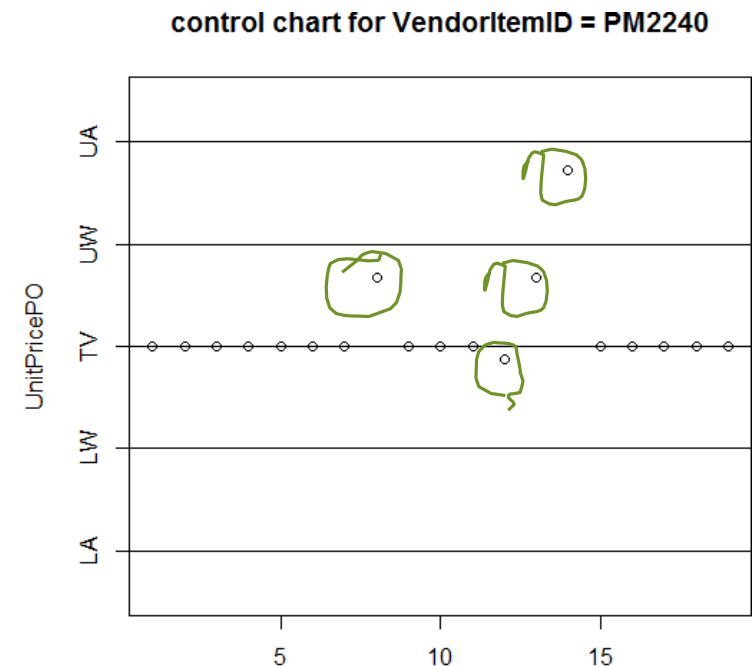
# Workout Sample

Understand your outliers = potential NC items

VendorName	VendorItemID	InvoiceDate	TAGItemDescr	InvoiceNumber	UnitPricePO	QuantityInv	LineProductAmount
ST JUDE MEDICAL SC INC	PM2240	3/16/2015	PM2240	9004644596	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/16/2015	PM2240	9004644597	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797306	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797312	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/29/2015	PM2240	9004838534	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/20/2015	PM2240	9004659359	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/20/2015	PM2240	9004735684	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/6/2015	PM2240	9004623708	4375	1	4375
ST JUDE MEDICAL SC INC	PM2240	5/26/2015	PM2240	9004826730	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/2/2015	PM2240	9004694016	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/4/2015	PM2240	9004702052	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/26/2015	PM2240	9004672671	3313.13	10	33131.3
ST JUDE MEDICAL SC INC	PM2240	3/3/2015	PM2240	9004610859	4375	1	4375
ST JUDE MEDICAL SC INC	PM2240	3/10/2015	PM2240	9004636604	5750	1	5750
ST JUDE MEDICAL SC INC	PM2240	3/25/2015	PM2240	9004668983	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/26/2015	PM2240	9004826734	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/23/2015	PM2240	9004746612	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/2/2015	PM2240	9004694014	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797313	3487.5	1	3487.5

4 items are clearly visible to be away from the target # as highlighted

VendorName = ST JUDE MEDICAL SC INC  
Contract effective from (2015-03-01) - (2017-01-31)  
VendorItemID = PM 2240



RESULT = PASS (15/19) FAIL (4/19)

# Workout Sample

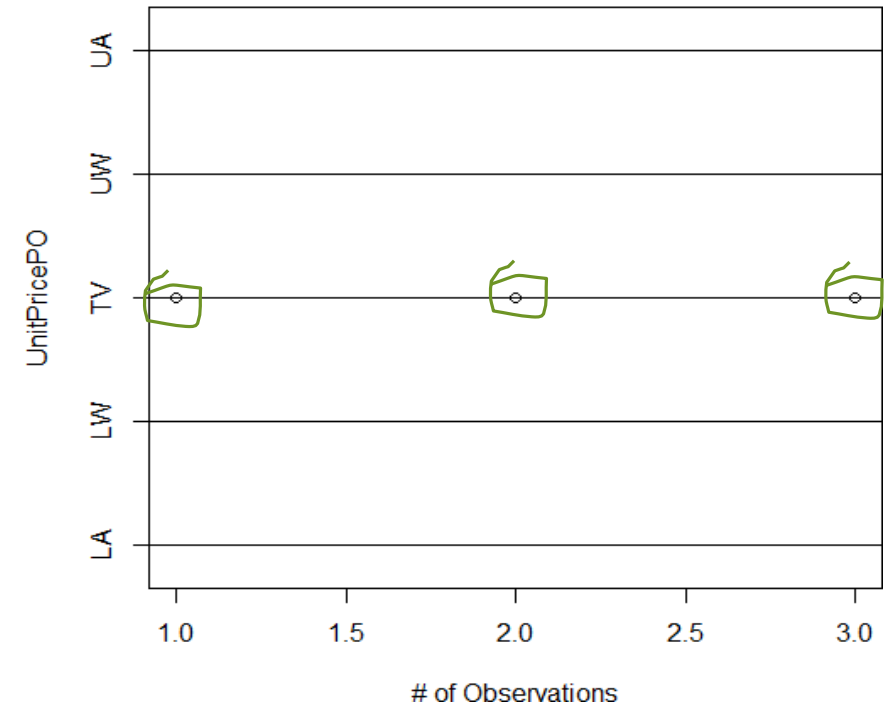
Understand your outliers = potential NC items

VendorName = ST JUDE MEDICAL SC INC  
Contract effective from (2015-03-01) - (2017-01-31)  
VendorItemID = PM 2160

VendorName	VendorItemID	InvoiceDate	TAGItemDescr	InvoiceNumber	UnitPricePO	QuantityInv	LineProductAmount
ST JUDE MEDICAL SC INC	PM2160	5/13/2015	PM2160	9004797310	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2160	4/4/2015	PM2160	9004702051	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2160	5/13/2015	PM2160	9004797311	3487.5	1	3487.5

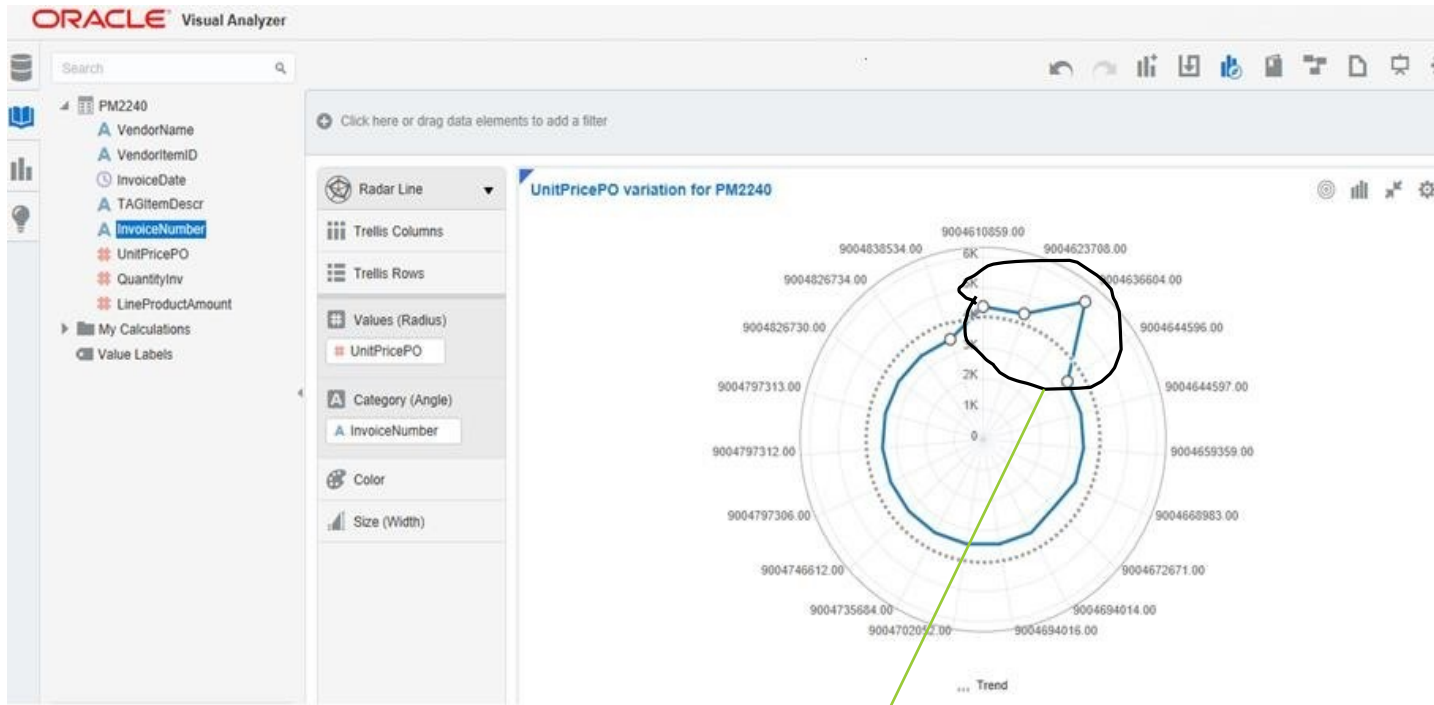
RESULT = PASS (3/3)

control chart for VendorItemID = PM2160





# Visualization of Contract/Invoice Matching



RESULT = 4 outlier observations

Table 1

Contract Number	Effective Date	Expiration Date	Term of Agreement (months)	Invoice Reference Number	Contract Number	Invoice Raised Date	Invoice Total Amount
CCG STJ	March 1, 2015	January 31, 2017	23 Months	INV012671	CCG STJ	October 10, 2016	\$2,754.17

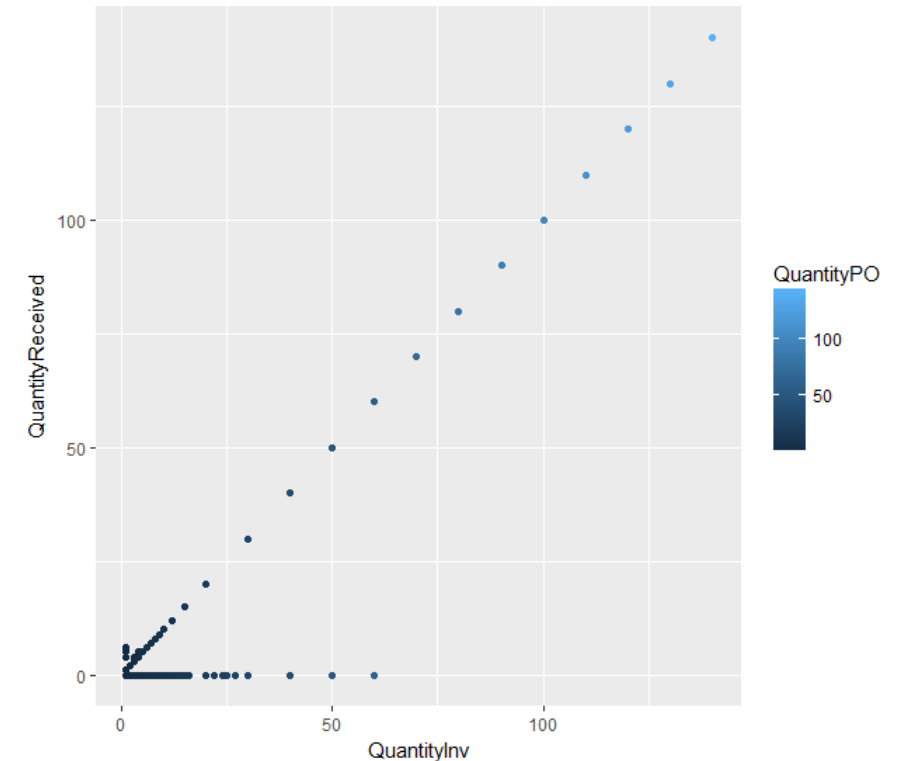
Matching of Contract # from Contract Doc as well from Invoice Doc

# Data Mining from Contracts and Invoices

Derive Combined Insights from your enterprise documents

APLineID	TAGItemC	VendorIt	VendorIt	ClientIt	ClientIt	UnitPrice	UnitPrice	QuantityReceiv	QuantityPO	QuantityInv	UOM	LineProdu	INVID	InvoiceNu	Schedulef	InvoiceDa	PONumbe	POLine
2.65E+08	401904	401904	CATH ABL	99400	CATH ABL	995	995	2	2	2	EA	1990	1.51E+08	9E+09	0	#####	2E+08	1
2.65E+08	2088TC58	2088TC/58	LEAD PACI	112989	LEAD PACI	400	400	1	1	1	EA	400	1.51E+08	9E+09	0	#####	4E+08	2
2.65E+08	DS2C014	DS2C014	INTRO SHI	114356	INTRO SHI	400	400	1	1	1	EA	400	1.52E+08	9E+09	0	#####	2E+08	4
2.65E+08	27AGFN75	27AGFN-7	VALVE AO	35883	VALVE AO	4516	4516	1	1	1	EA	4516	1.51E+08	9E+09	0	#####	2E+08	1
2.65E+08	2434369	243-4369	PROBE DN	136563	PROBE DN	243.11	243.11	0	1	1	EA	243.11	1.51E+08	9.06E+08	0	#####	1.5E+08	2
2.65E+08	127240	127240	KIT ABLAT	127240	KIT ABLAT	837	837	0	25	25	EA	20925	1.51E+08	9E+09	0	#####	2E+08	1
2.65E+08	135440	135440	AB CD15 C	135440	AB CD15 C	206.24	206.24	1	1	1	EA	206.24	1.51E+08	9.05E+08	0	3/5/2014	1.5E+08	1
2.65E+08	4.93E+09	4.93E+09	TUBE ANL	54581	TUBE ANL	229.02	229.02	6	6	6	PK	1374.12	1.51E+08	9.05E+08	0	4/9/2014	4E+08	1
2.65E+08	EX3151	EX3151	CABLE CO	127920	CABLE CO	83	83	0	1	1	EA	83	1.51E+08	9E+09	0	#####	2E+08	7
2.65E+08	760700	760-700	KIT IMMTS	136515	KIT IMMTS	1656.51	1656.51	0	2	2	EA	3313.02	1.51E+08	9.06E+08	0	4/1/2015	1.5E+08	1
2.65E+08	2088TC52	2088TC/52	LEAD PACI	113511	LEAD PACI	400	400	1	1	1	EA	400	1.51E+08	9E+09	0	#####	2E+08	3
2.65E+08	407206	407206	NDL 98CM	114978	NDL 98CM	210	210	0	1	1	EA	210	1.51E+08	9E+09	0	#####	2E+08	1
2.65E+08	1.19E+10	1.19E+10	CHEMSTRI	X-BC-200743		28.57	28.57	6	6	6	EA	171.42	1.52E+08	9.05E+08	0	#####	1.5E+08	1
2.65E+08	3003248	3003248	RGT PCR H	80999	RGT PCR H	183.6	183.6	20	20	20	EA	3672	1.51E+08	9.04E+08	0	#####	4E+08	1
2.65E+08	401764	401764	CATH EP 1	123302	CATH EP 1	110	110	2	2	2	EA	220	1.51E+08	9E+09	0	#####	1.5E+08	1
2.65E+08	406840	406840	INTRO SHI	16635	INTRO SHI	165	165	0	10	10	EA	1650	1.51E+08	9E+09	0	2/5/2015	1.5E+08	2
2.65E+08	BVCS6180	BVCS6180	CATH ANG	126436	CATH ANG	100	100	0	1	1	EA	100	1.51E+08	9E+09	0	#####	2E+08	5
2.65E+08	406849	406849	INTRO SHI	60848	INTRO SHI	165	165	0	1	1	EA	165	1.51E+08	9E+09	0	5/8/2015	2E+08	4
2.65E+08	405520	405520	DIL VASC	95514	DIL VASC	4.2	4.2	10	10	10	EA	42	1.51E+08	9E+09	0	#####	4E+08	1
2.65E+08	5.24E+09	5.24E+09	KIT SAMPL	147009	KIT SAMPL	625	625	0	1	1	KT	625	1.51E+08	9.06E+08	0	#####	4E+08	2
2.65E+08	406501	406501	SIV CATH	135242	SIV CATH	41.25	41.25	2	2	2	BX	82.5	1.51E+08	9E+09	0	1/6/2014	1.5E+08	1

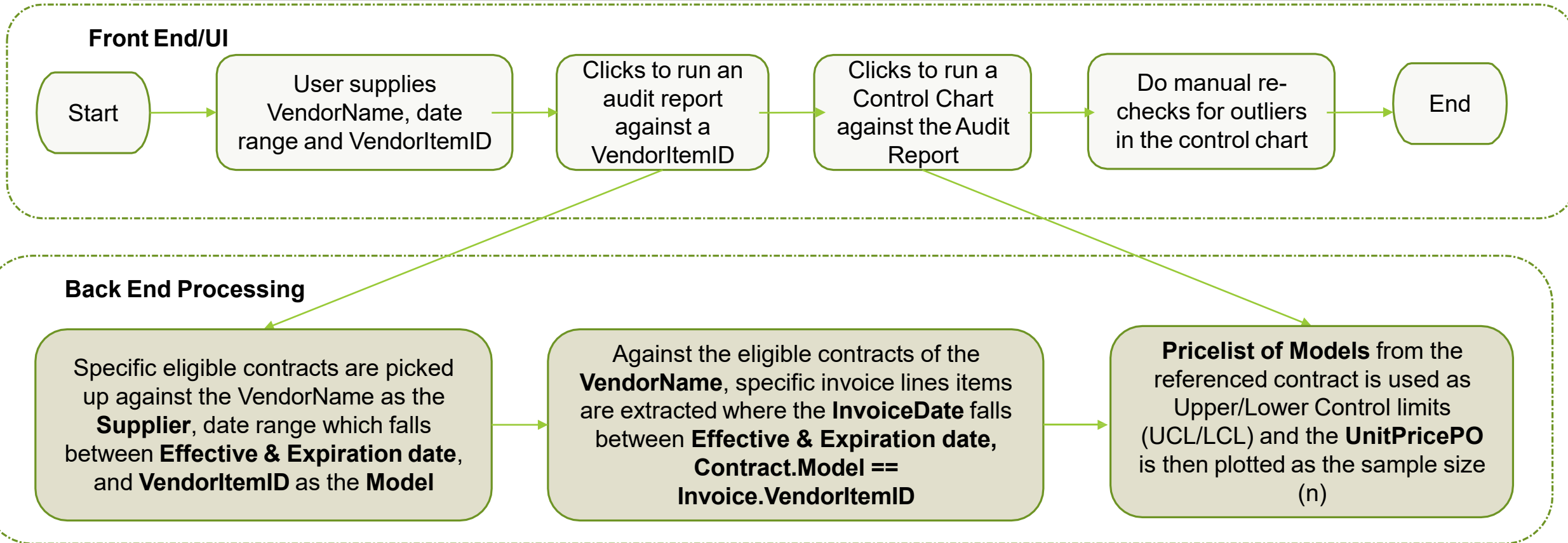
Items highlighted in red – there is a mismatch of QtyReceived vs Invoiced or Requested/PO



*Corresponding items not yet acknowledged by the material receipts manager*

# Sample Workout

# Functional Flow



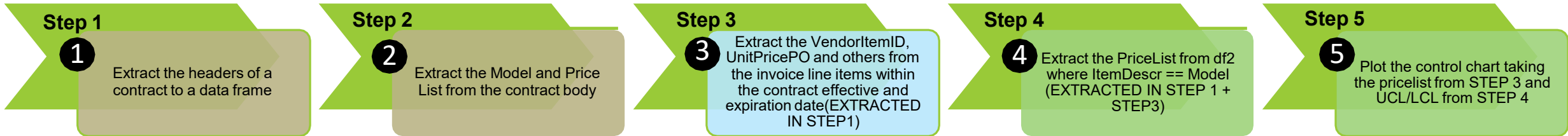
# Sample Workout

# Technical Flow

Contract Invoice Derived

```
4 > df22$f1 > myvector = df_sjude3$UnitPricePO
  [1] PM2240 > myvector
  [2] PM2240 [1] 3487.50 3487.50 3487.50 3487.50 3487.50 3487.50 3487.50 4375.00 3487.50
  [3] 3875.5 [10] 3487.50 3487.50 3313.13 4375.00 5750.00 3487.50 3487.50 3487.50 3487.50
  [4] 3487.5 [19] 3487.50
  [5] 2950
```

To act as UCL and LCL for our control chart



a1	a2
Contract Number	CCG ST]
Effective Date	March 1, 2015
Expiration Date	January 31, 2017
Term of Agreement (months)	23 Months

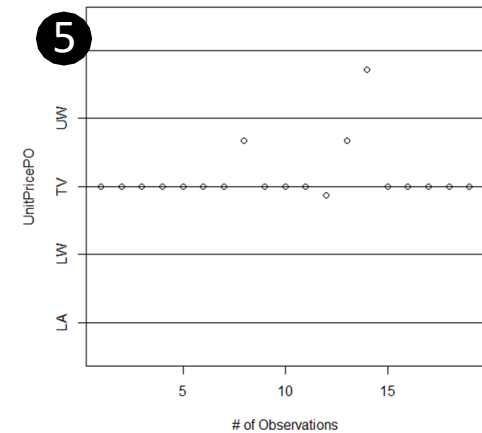
df1

f0	f1	f2	f3
Model	PM2240	PM2160	PM2210
Description	PM2240 Assurity Pacemaker DR - RF 3487.50 2950.00	PM2160 Endurity Pacemaker DR 3487.50 2950.00	PM2210 Accent DR RF 3487.50 2950.00
Access	3875.5	3758.75	3758.75
Dual 85%	3487.5	3487.5	3487.5
Dual 85% IDN	2950	2950	2950

df3

VendorName	VendorItemID	InvoiceDate	TAGItemDescr	InvoiceNumber	UnitPricePO	QuantityInv	LineProductAmount
ST JUDE MEDICAL SC INC	PM2240	3/16/2015	PM2240	9004644596	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/16/2015	PM2240	9004644597	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797306	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797312	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/29/2015	PM2240	9004838534	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/20/2015	PM2240	9004659359	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/20/2015	PM2240	9004735684	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/6/2015	PM2240	9004623708	4375	1	4375
ST JUDE MEDICAL SC INC	PM2240	5/26/2015	PM2240	9004826730	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/2/2015	PM2240	9004694016	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/4/2015	PM2240	9004702052	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/26/2015	PM2240	9004672671	3313.13	10	3313.13
ST JUDE MEDICAL SC INC	PM2240	3/3/2015	PM2240	9004610859	4375	1	4375
ST JUDE MEDICAL SC INC	PM2240	3/10/2015	PM2240	9004636604	5750	1	5750
ST JUDE MEDICAL SC INC	PM2240	3/25/2015	PM2240	9004668983	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/26/2015	PM2240	9004826734	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/23/2015	PM2240	9004746612	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/2/2015	PM2240	9004694014	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797313	3487.5	1	3487.5

control chart for TAGItemDescr = PM2240



# Sample Workout

## Snapshots from Contracts and Invoice lines

**IDN\_ ST. JUDE MEDICAL CARDIAC RHYTHM MANAGEMENT AGREEMENT  
COVER SHEET**

CONTRACT NUMBER: CCG STJ

1. The "Parties" to this Agreement are  
**PeaceHealth**  
("Account")

**St. Jude Medical S.C., Inc.**  
("Seller")

1115 SE 164<sup>th</sup>  
Vancouver, WA 98683  
Attention: Stacy Roberts  
Contract Manager  
(Phone) 360-729-1803

6300 Bee Cave Rd.  
Bldg. 2, Ste. 100, Austin, TX 78746  
Attention: USD Contracts  
512.286.4000 (Phone) 818.256.8330 (Fax)

2. Product Category: **Cardiac Rhythm Management ("CRM") Devices**  
3. Effective Date: **March 1, 2015**  
4. Expiration Date: **January 31, 2017**  
5. Term of Agreement (months): **23 Months**  
6. Guaranteed Delivery Time Period (days): **3** (Section 6.5)  
7. Commitment: **Vendor All Play Multi-Source Award or Dual Vendor 85%**

### Contract Header

Dates from contract header and Invoice lines are matched, then itemlist and pricelist are matched.

APLineID	TACline	Vendo	Vendo	Client	Client	UnitPri	UnitPri	Quanti	Quanti	Quanti	UOM	LinePri	NVID	InvoiceNumber	Sched	InvoiceDt	PDNuri	POLink	SeqNE	Entty	Vendo	VendorName
265417440	PM2240	PM2240	PACERC	147766	PACERC	4375	4375	0	1	1	EA	4375	151472757	9004196908	0	3/15/2014	2E+08	3	3	200	103759	ST JUDE MEDICAL SC INC
265418567	PM2240	PM2240	PACERC	147766	PACERC	4375	4375	0	1	1	EA	4375	151549713	9004397809	0	12/14/2014	4E+08	3	3	400	103759	ST JUDE MEDICAL SC INC
265424487	PM2240	PM2240	PACERC	147766	PACERC	4375	4375	0	1	1	EA	4375	151040635	9004378890	0	11/26/2014	2E+08	3	3	200	103759	ST JUDE MEDICAL SC INC
265424553	PM2240	PM2240	PACERC	147766	PACERC	4375	4375	0	1	1	EA	4375	151802230	9004453553	0	12/29/2014	2E+08	4	4	200	103759	ST JUDE MEDICAL SC INC
265425559	PM2240	PM2240	PACERC	147766	PACERC	3487.5	3487.5	0	1	1	EA	3487.5	150892666	9004644596	0	3/16/2015	4E+08	1	1	400	103759	ST JUDE MEDICAL SC INC
265420856	PM2240	PM2240	PACERC	147766	PACERC	3487.5	3487.5	0	1	1	EA	3487.5	151322731	9004644597	0	3/16/2015	4E+08	1	1	400	103759	ST JUDE MEDICAL SC INC
265430935	PM2240	PM2240	PACERC	147766	PACERC	3487.5	3487.5	0	1	1	EA	3487.5	151323432	9004797306	0	5/13/2015	4E+08	1	1	400	103759	ST JUDE MEDICAL SC INC
265431261	PM2240	PM2240	PACERC	147766	PACERC	4156.25	4156.25	0	10	10	EA	4156.25	151247408	9004488536	0	1/7/2015	2E+08	4	4	200	105934	ST JUDE MEDICAL CRDM
265455147	PM2240	PM2240	PACERC	147766	PACERC	3487.5	3487.5	0	1	1	EA	3487.5	150950670	9004797312	0	5/13/2015	4E+08	1	1	400	103759	ST JUDE MEDICAL SC INC
265457805	PM2240	PM2240	PACERC	147766	PACERC	3487.5	3487.5	0	1	1	EA	3487.5	151172758	9004838534	0	5/29/2015	4E+08	1	1	400	103759	ST JUDE MEDICAL SC INC

### Invoice Lines

#### EXHIBIT A PRODUCTS AND PRICING

All accessories, with the exception of LV lead delivery systems\*, are included at no additional cost with a St. Jude Medical device implant. \*\*Products included in LV lead delivery system are detailed herein

Model	Description	Access	Dual 85% Facility Level	Dual 85% IDN Level
<b>Brady Pacing Products</b>				
<b>Dual Chamber DDR Pacemakers</b>				
PM2240	Assurity Pacemaker DR - RF		3,875.00	3,487.50
PM2160	Endurity Pacemaker DR		3,758.75	3,487.50
PM2210	Accent DR RF		3,758.75	3,487.50
PM2110	Accent DR		3,758.75	3,487.50
5826	Zephyr XL DR		3,007.00	2,790.00
5820	Zephyr DR		3,007.00	2,790.00
5357M/S	Verity ADx XL DR M/S		3,007.00	2,790.00
<b>Single Chamber SSIR Pacemakers</b>				
PM1240	Assurity Pacemaker SR - RF		3,200.00	2,880.00
PM1160	Endurity Pacemaker SR		3,104.00	2,880.00
PM1210	Accent SR RF		3,104.00	2,880.00

### Contract Body

# Sample Workout

## Typical Wireframe (1/2)



Step 1

Select a Vendor Name	ST JUDE MEDICAL SC INC
	ST JUDE MEDICAL SC INC
Search Invoices (From Date)	ST JUDE MEDICAL SC INC
Search Invoices (To Date)	ST JUDE MEDICAL CARDIAC SU
	ROCHE DIAGNOSTICS CORP
	ST JUDE MEDICAL SC INC
Select a VendorItemID	ROCHE DIAGNOSTICS CORP
	ROCHE DIAGNOSTICS CORP
	Run an Audit Report

Step 3

Select a Vendor Name	ST JUDE MEDICAL SC INC
Search Invoices (From Date)	3/1/2015
Search Invoices (To Date)	1/31/2017
Select a VendorItemID	5883482001
	85785
	5883466001
	406898
	760-2037
	139791
	139051
	PM2240
	100042049

Select a Vendor Name	ST JUDE MEDICAL SC INC
Search Invoices (From Date)	3/1/2015
Search Invoices (To Date)	1/31/2017
Select a VendorItemID	5883482001
	Run an Audit Report

Step 2

Select a Vendor Name	ST JUDE MEDICAL SC INC
Search Invoices (From Date)	3/1/2015
Search Invoices (To Date)	1/31/2017
Select a VendorItemID	PM2240
	Run an Audit Report

Step 4

# Sample Workout

## Typical Wireframe

(2/2)

Step 5

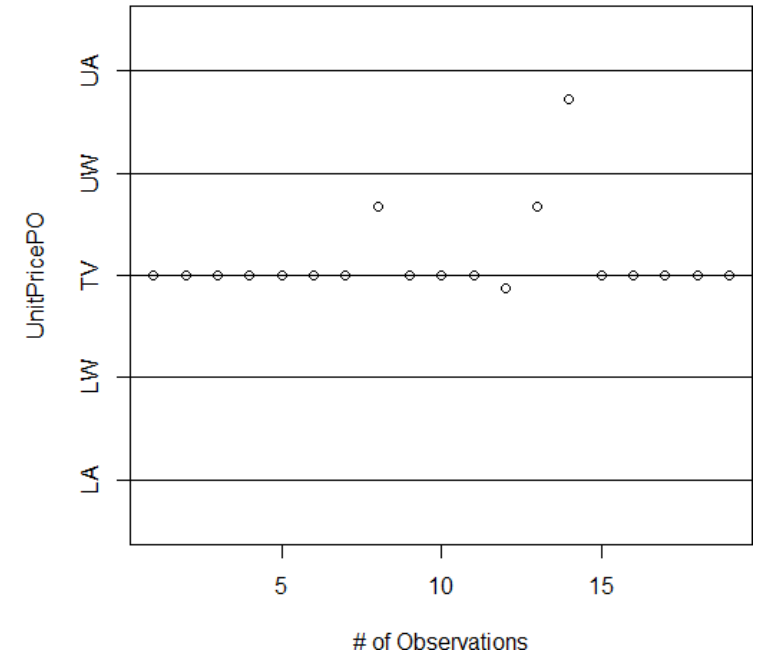


VendorName	VendorItemID	InvoiceDate	TAGItemDescr	InvoiceNumber	UnitPricePO	QuantityInv	LineProductAmount
ST JUDE MEDICAL SC INC	PM2240	3/16/2015	PM2240	9004644596	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/16/2015	PM2240	9004644597	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797306	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797312	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/29/2015	PM2240	9004838534	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/20/2015	PM2240	9004659359	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/20/2015	PM2240	9004735684	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/6/2015	PM2240	9004623708	4375	1	4375
ST JUDE MEDICAL SC INC	PM2240	5/26/2015	PM2240	9004826730	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/2/2015	PM2240	9004694016	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/4/2015	PM2240	9004702052	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	3/26/2015	PM2240	9004672671	3313.13	10	33131.3
ST JUDE MEDICAL SC INC	PM2240	3/3/2015	PM2240	9004610859	4375	1	4375
ST JUDE MEDICAL SC INC	PM2240	3/10/2015	PM2240	9004636604	5750	1	5750
ST JUDE MEDICAL SC INC	PM2240	3/25/2015	PM2240	9004668983	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/26/2015	PM2240	9004826734	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/23/2015	PM2240	9004746612	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	4/2/2015	PM2240	9004694014	3487.5	1	3487.5
ST JUDE MEDICAL SC INC	PM2240	5/13/2015	PM2240	9004797313	3487.5	1	3487.5

Run the Control Chart

Step 6

control chart for VendorItemID = PM2240



RESULT = PASS (15/19) FAIL (4/19)



# Thank You

Greenojō provides Automation, Analytics and AI solutions to enterprise customers

---

For RFPs, Solutions and Sales/Partner enquiries, connect us at - [sales@greenojō.com](mailto:sales@greenojō.com)