

14 Kasım 2012 - Rixos Ankara Oteli



# IBM Connected 2012 Ankara

## Yüksek Performanslı Veri Ambarı ve Analitik

Ayhan Önder

Pure Data for Analytics CTP



# IBM Pure Data System for Analytics

Powered by  
Netezza Technology

*Analitik İhtiyaçlarınız için İş Yükü Optimize Sistem*

## Analitik Uygulamalar

BI / Raporlama | Analiz/ Görselleştirme | Fonksiyonel Uyg. | Endüstri Uyg. | Veri Madenciliği | İçerik Analitiği

## IBM Big Data Platformu

Görselleştirme  
ve Keşif

Uygulama  
Geliştirme

Sistem  
Yönetimi



Hızlandırıcılar

Hadoop  
Sistemi



Akışkan  
veri işleme



Veri  
Ambarı



Veri Entegrasyonu ve Sahipliği

PureData  
System for Analytics





# PureData System for Analytics



## Kullanıcı deneyimini bütünüyle değiştiriyor

- ✓ Amaç optimize Analitik sistem
- ✓ Entegre Veri Tabanı, Sunucu ve Veri Saklama Ünitesi
- ✓ Standart Arayüzler üzerinden erişim
- ✓ Düşük Sahip Olma Maliyeti

**Hız:** Geleneksel çözümlere\* göre 10-100 kat hızlı

**Basitlik:** Minimal yönetim ve optimizasyon

**Ölçeklenebilirlik:** Petabyte ölçeğinde veri kapasitesi

**Akıllı:** Yüksek performanslı ileri düzey analitik

\* IBM Müşterilerden alınan geri bildirimlere dayanmaktadır. "Geleneksel çözümler" ifadesiyle profesyonel olarak bütün olarak imal edilmiş iş yükü optimize sistemler dışındaki sistemler kastedilmektedir.

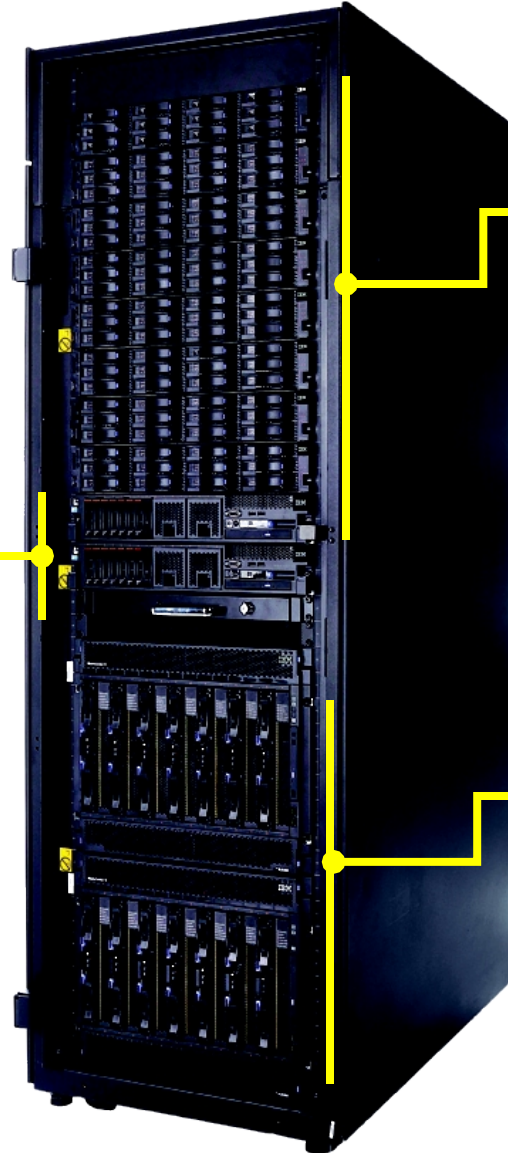
# “IBM PureData System for Analytics” içerisinde neler var?

## Optimize Donanım ve Yazılım

- Donanım Hızlandırıcılı AMPP Mimari
- Yüksek performanslı analitik amaçlı tasarım
- Optimizasyon ihtiyacı yok

## SMP Host

- SQL Çözümleyici
  - Sorgu Planları
- Otomatik Optimizasyon
  - Yönetim



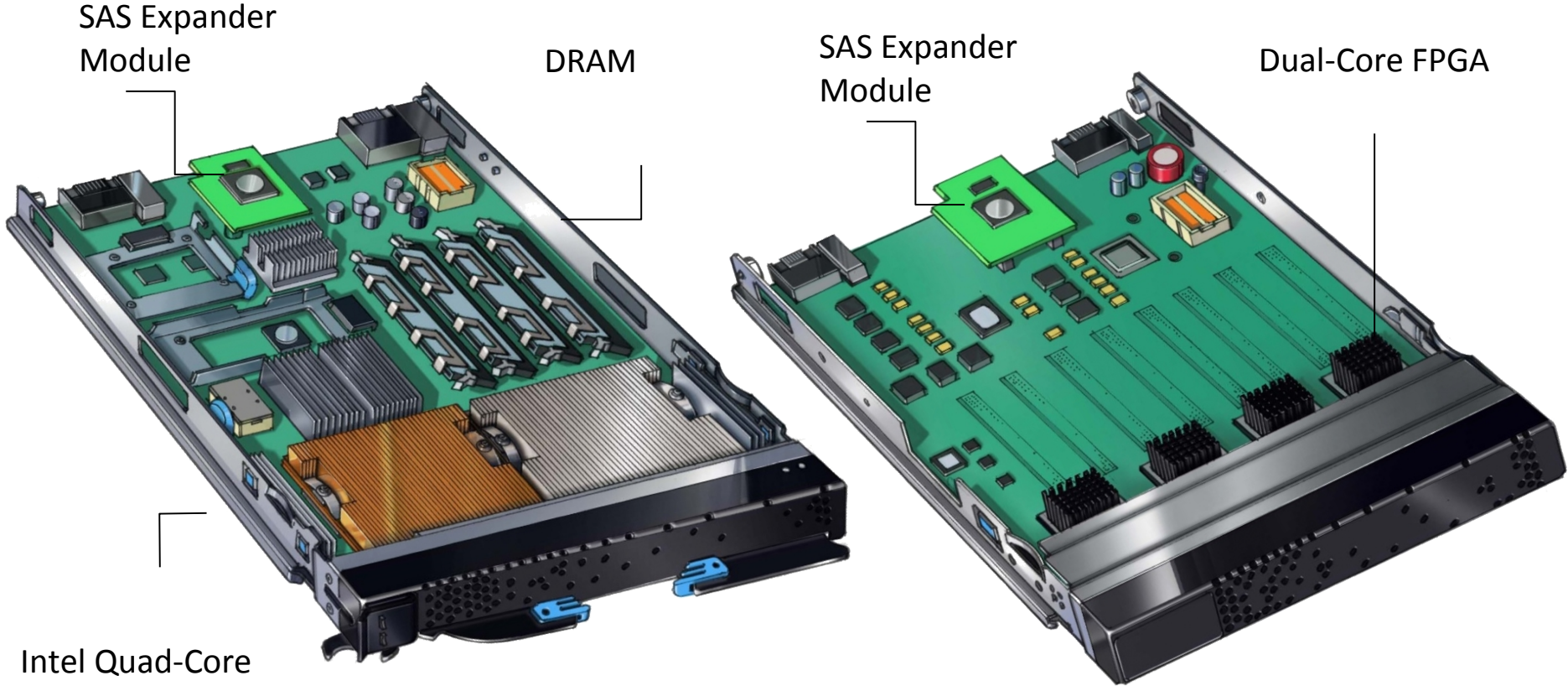
## Disk Ünitesi

- Kullanıcı verileri, aynalama, swap alanı
- Yüksek performanslı veri tarama

## Snippet Blades™

- FPGA bazlı donanım temelli sorgu hızlandırıcı
- Göz alıcı yüksek performans
- Kompleks analitik fonksiyonlar veriler diskten okunurken işlenir

## Snippet-Blade™ (S-Blade) Bileşenleri

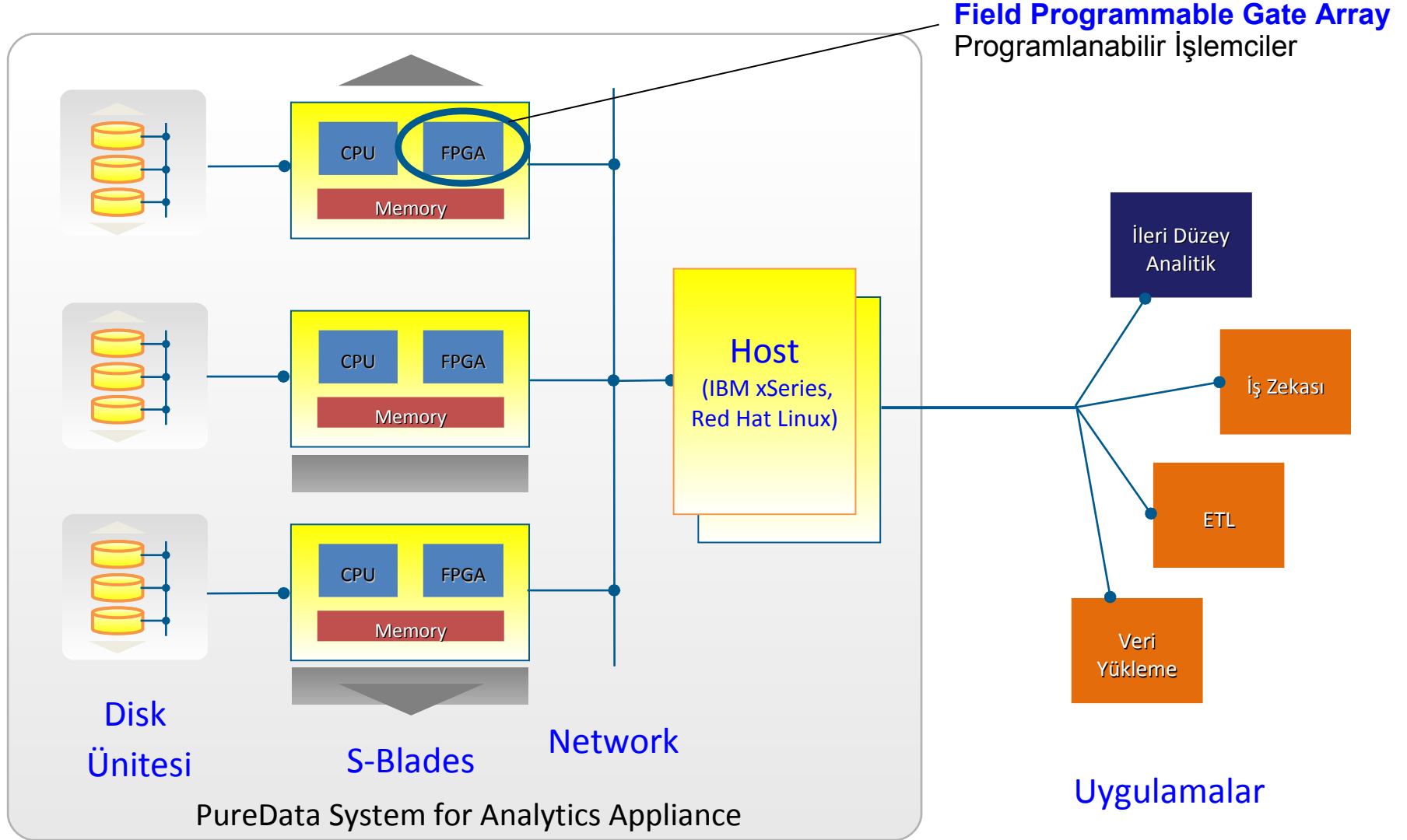


IBM BladeCenter Sunucu

Netezza Veri Tabanı Hızlandırıcısı



## PureData System for Analytics, AMPP Mimarisi



## PureData System for Analytics'i farklı kılan nedir?



HIZ

Eskisine nazaran **2000 kat** hızlı  
Her ay **%30** büyüme

*"Netezza has allowed us to reduce the complexity of regulatory reporting and processing of exchange data from days down to minutes."*



BASİTLİK

Herhangi bir eğitim almadan **6 ay** boyunca üretim ortamında Oracle sistemine göre **200 kat** Daha hızlı, **3 ayda** yatırımın geri dönüşü

*"Allowing the business users access to the Netezza box was what sold it."*

- Steve Taff, Executive Dir. of IT Services



ÖLÇEKLENEBİLİRLİK

**7 yıllık tarihçe**  
**1 PB**'lık veri Netezza'da saklanıyor  
**%100-200** yıllık büyüme oranı

*"NYSE ... has replaced an Oracle IO relational database with a data warehousing appliance from Netezza, allowing it to conduct rapid searches of 650 terabytes of data."*



AKILLI

SUNY Buffalo araştırmacıları **kentilyonlarca hesaplamayı** **27 saatten 12 dakikaya** indirdi

*"Once we had the data on Netezza we were able to do the same analysis and much more complex analysis in minutes. The research draws on medical records, lab results, MRI scans, and patient surveys."*

- Dr. Murali Ramanathan, SUNY Buffalo







# Entegre Analitik

IBM Netezza In-Database Analytics Version 2.0

## Netezza Veri Tabanı İçi Analitik

Transformasyonlar

Matematik

Coğrafi Analiz

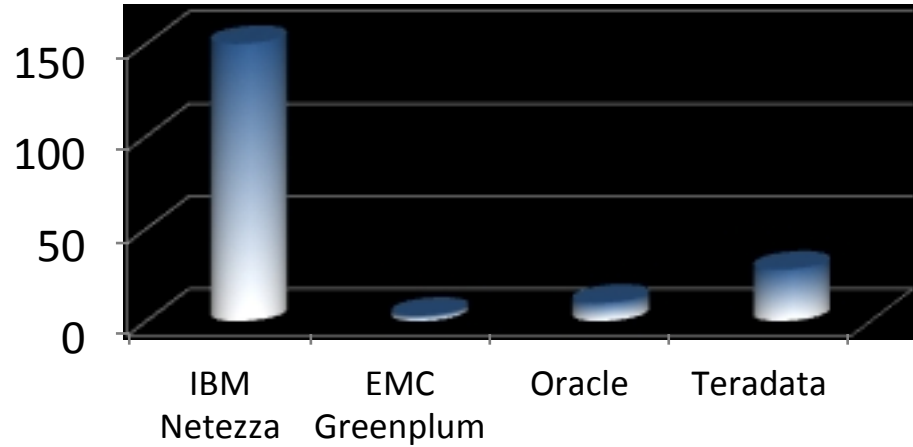
Tahminleme

İstatistik

Zaman Serileri

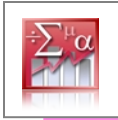
Veri Madenciliği

## EN GENİŞ Analitik Kütüphane



- Veriyi yerinden çıkarmadan
- Derinlemesine ve büyük veri analizi
- Yüksek performanslı, paralel hesaplama

# Hazır Veri Tabanı içi Analitik Yetkinlikler



## Statistics

- Descriptive Statistics+
- Distance Measures\*
- Hypothesis Testing\*
- Chi-Square & Contingency Tables\*
- Univariate & Multivariate Distributions+
- Monte Carlo Simulation\*



## Transformation

- Data Profiling / Descriptive Statistics+
- General Diagnostics
- Statistics+
- Sampling
- Data prep



## Time Series

- Autoregressive+
- Forecasting\*



## Mathematical

- Basic Math\*
- Permutation and Combination\*
- Greatest Common Divisor and Least Common Multiple\*
- Conversion of Values\*
- Exponential and Logarithm\*
- Gamma and Beta Functions
- Matrix Algebra+
- Area Under Curve\*
- Interpolation Methods\*



## Data Mining

- Association Rules+
- Clustering+
- Feature Extraction+
- Discriminant Analysis\*



## Predictive

- Linear Regression+
- Logistic Regression+
- Classification
- Bayesian
- Sampling
- Model Testing



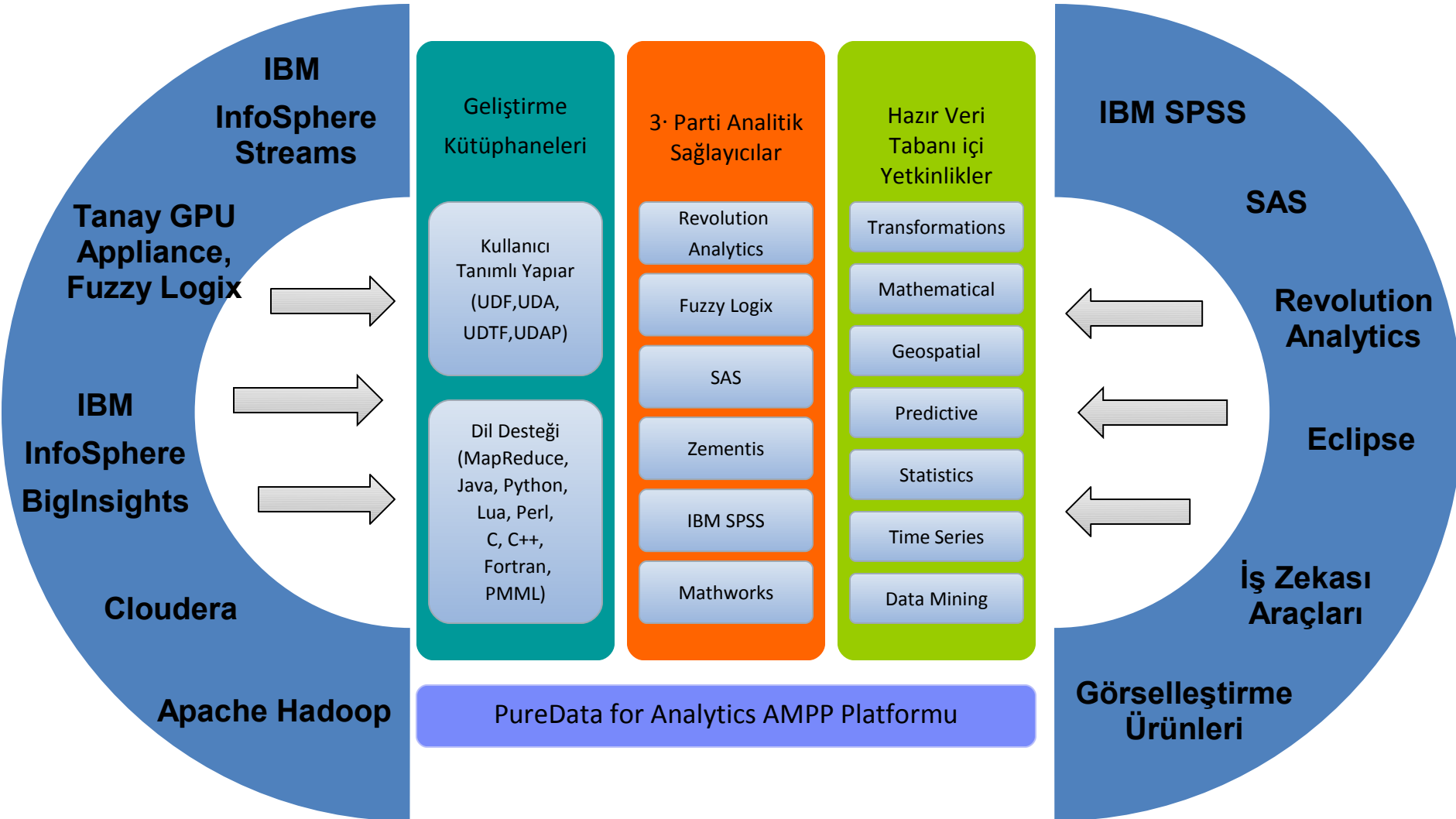
## Geospatial

- Geospatial Data Type
- Geometric Functions
- Geometric Analysis

\* Fuzzy Logix DB  
Lytix yetkinlikleri

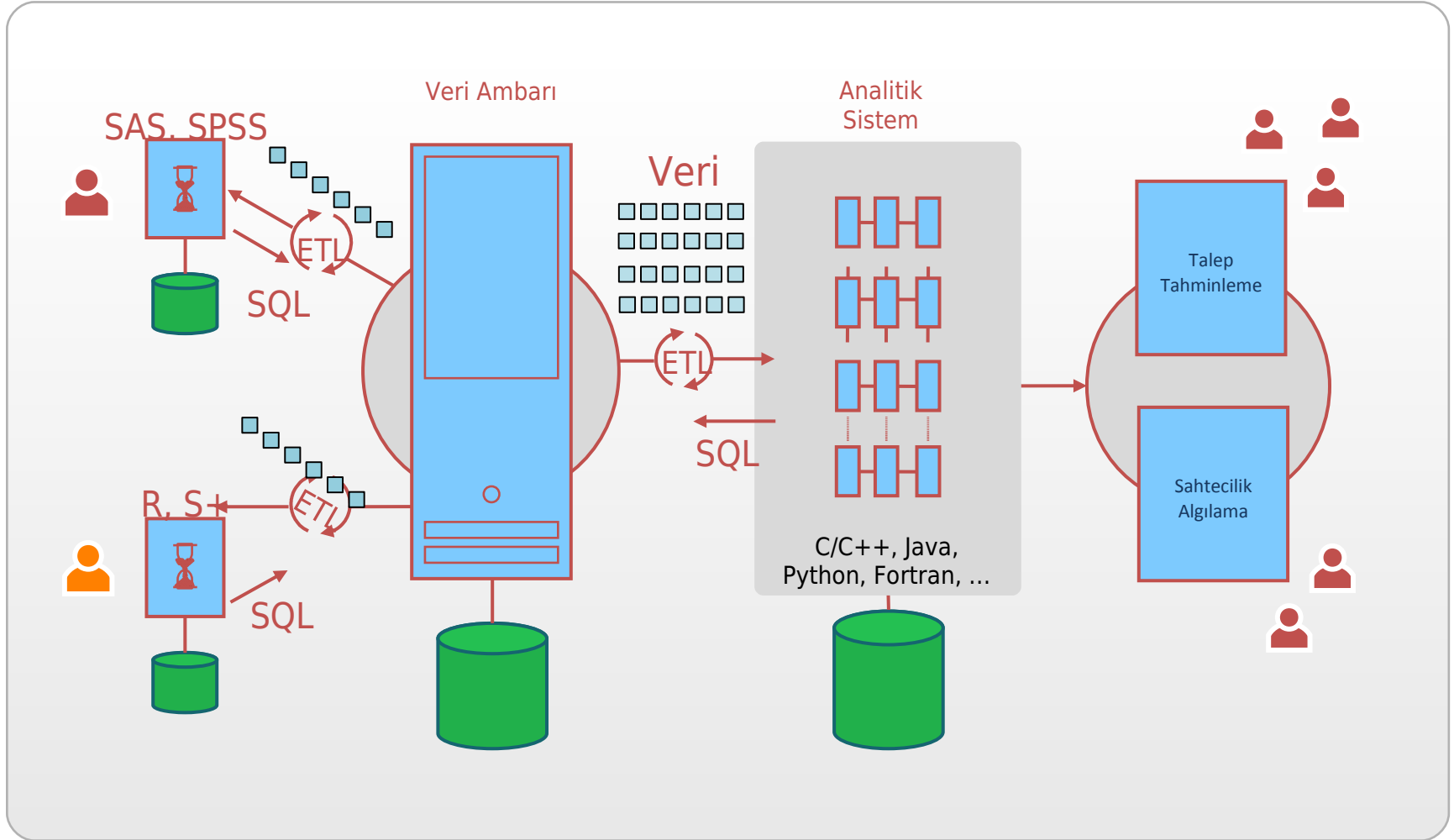
+ Netezza  
Analytics ve Fuzzy  
Logix DB Lytix  
yetkinlikleri

# IBM Netezza Analytics

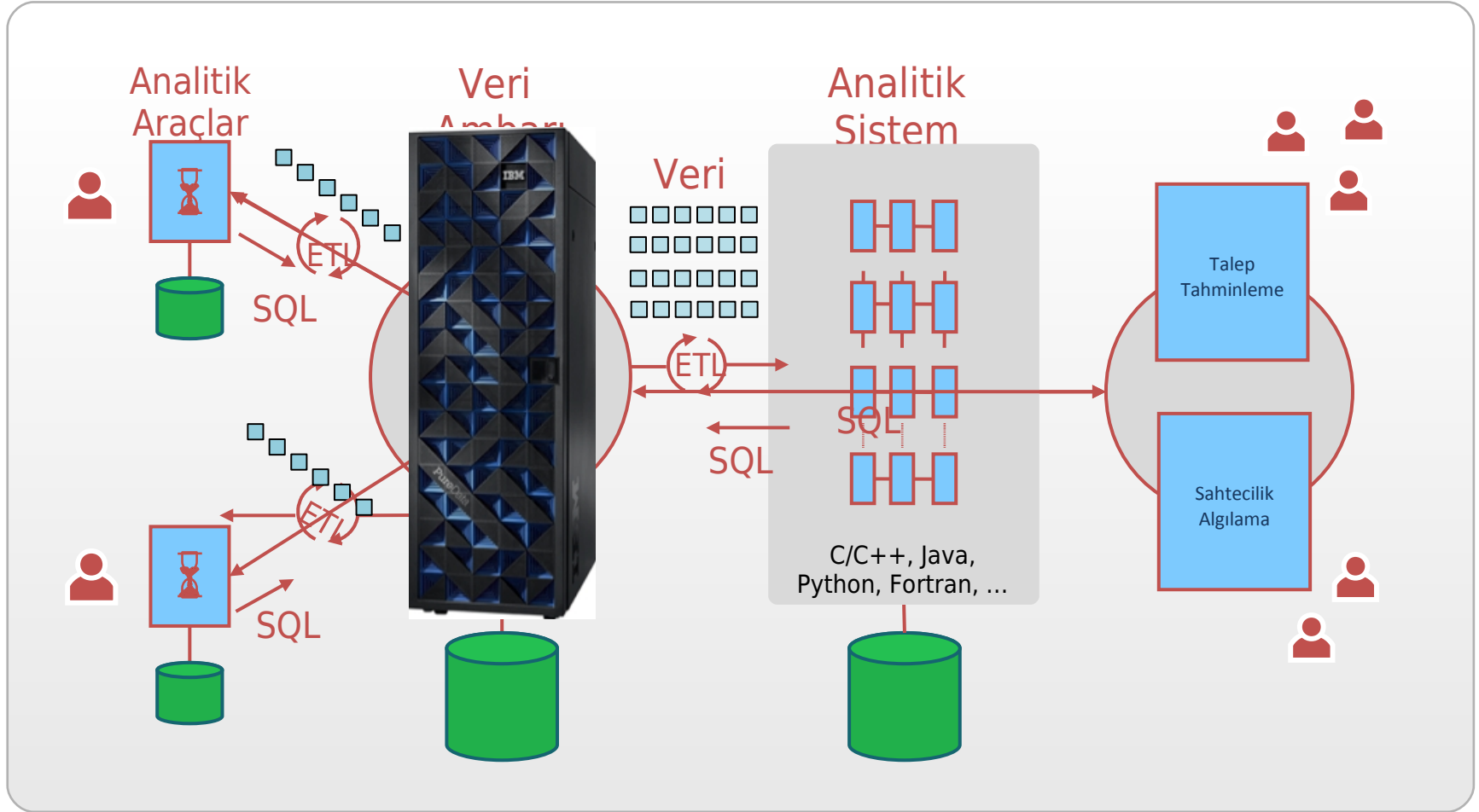




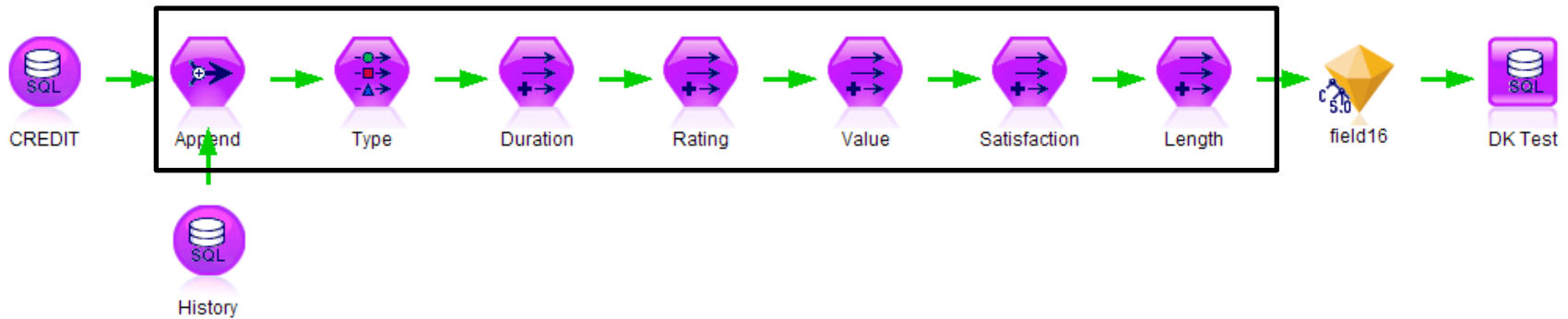
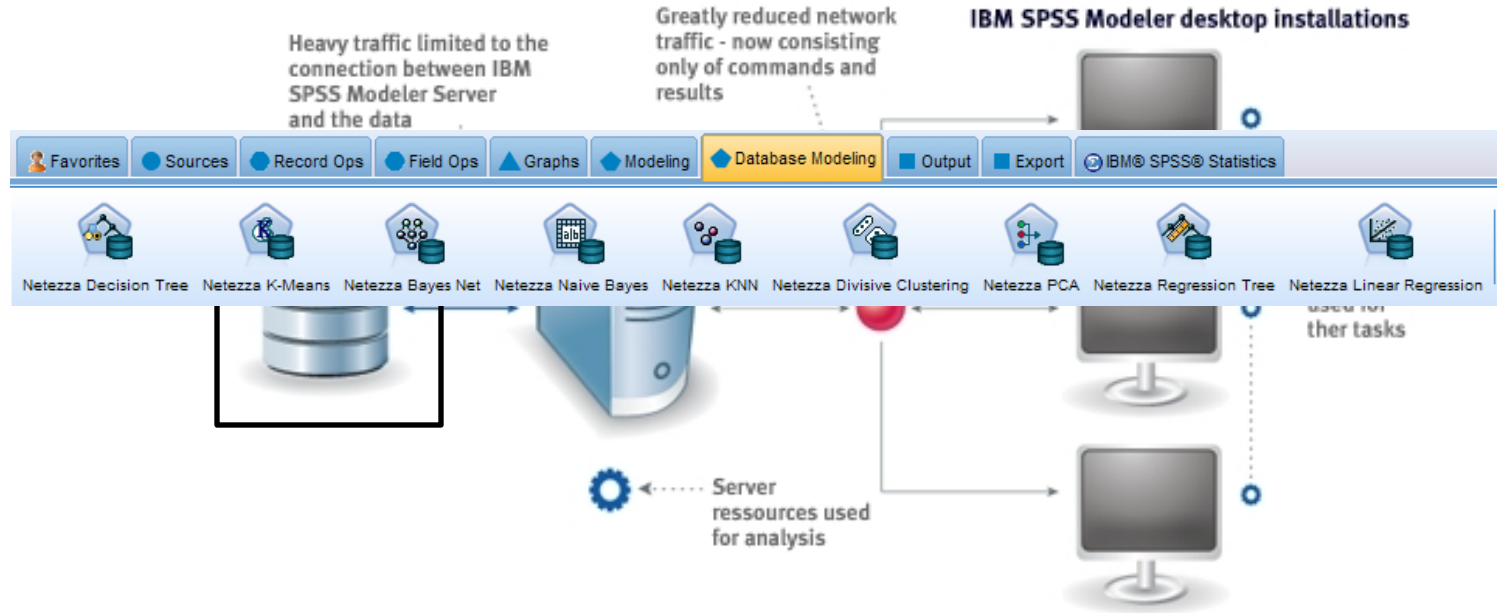
# İleri Düzey Analitik – Geleneksel Yöntem



# IBM Netezza Analytics – Zorlu Analitik Çalışmaları Basitleştirir

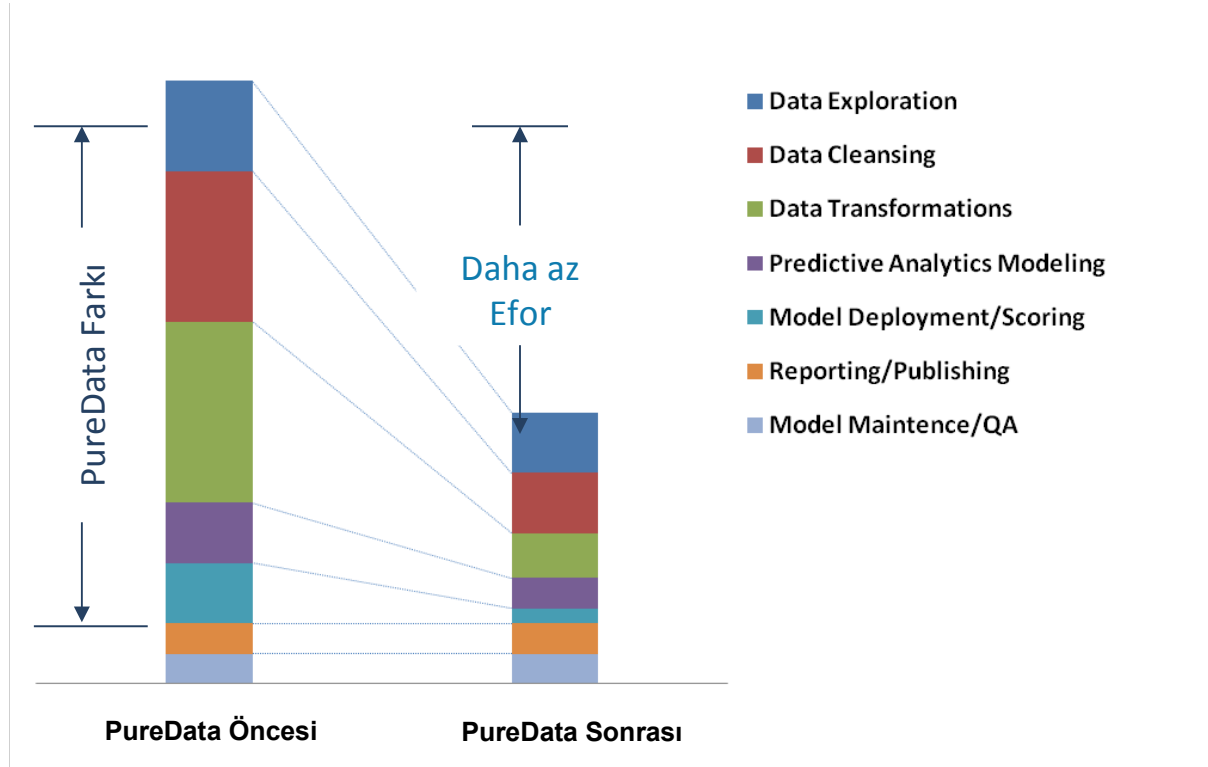


# IBM SPSS Modeler & IBM Pure Data for Analytics





## Veri Tabanı içi Analitik ile Verimliliğinizi arttırın



- ✓ **Basit – Verileri taşımayın**
- ✓ **Tutarlı – Örneklem kullanmak zorunda kalmayın**
- ✓ **Hızlı – Veri Tabanı içi skoreleme**

- ✓ Kolay sahip olunabilir
- ✓ Düşük altyapı maliyeti
- ✓ Yüksek Analist verimliliği