

Bab 5: Pengumpulan dan Analisis Data



Source: Harrell, C., B.K. Ghosh and R.O. Bowden, Jr., *Simulation Using Promodel*, 2nd ed., McGraw-Hill, Singapore, 2003.

Bab 5: Pengumpulan dan Analisis Data

2

Bacaan:

- Harrell, Bab 6
- www.teknikindustri.org

Topik

- Pengumpulan data:
 - Panduan Pengumpulan Data
 - **Menentukan Kebutuhan Data**
 - Identifikasi Sumber Data
 - Mengumpulkan Data
 - Membuat Asumsi
- Distribusi Frekuensi:
 - Analisis Statistik Data Numerik
 - *Distribution Fitting*
 - Memilih sebuah distribusi bila data tidak tersedia
 - *Bounded vs. Boundless Distributions*
 - Pemodelan Probabilitas Diskrit menggunakan Distribusi Kontinu
 - **Dokumentasi dan Persetujuan Data**

1. Pendahuluan





Pendahuluan

4

Dalam simulasi, pengumpulan data merupakan pekerjaan yang paling menantang dan menghabiskan waktu

- Sistem baru: hanya tersedia estimasi dalam garis besar
- Sistem saat ini: data yang ada mentah dan tak terorganisir



Pendahuluan (lanjutan)

5



- Bagaimana **prosedur** pengumpulan data yang paling baik?
- **Tipe** data apa yang harus dikumpulkan?
- **Sumber**-sumber apa yang sebaiknya digunakan?
- Analisis **tipe** apakah yang harus dilakukan?
- Bagaimana memilih **distribusi probabilitas** yang tepat?
- Bagaimana data sebaiknya **didokumentasikan**?

2. Panduan Pengumpulan Data





Panduan

7

1. Identifikasikan kejadian yang memicu (*triggering events*)
2. Fokus hanya pada faktor kunci yang berpengaruh
3. Isolasikan waktu aktivitas aktual
4. Carilah pengelompokan yang biasa digunakan
5. Fokus pada esensi, bukan pada substansi
6. Pisahkan variabel input dari variabel respon

1. Identifikasi kejadian yang memicu

8

- Apakah yang memicu pergerakan entitas?
- Apakah yang memicu sumber-sumber tertentu?

2. Fokus hanya pada faktor kunci yang berpengaruh

9

- Jangan menghabiskan waktu dengan meneliti informasi yang memiliki pengaruh kecil atau tidak berpengaruh terhadap kinerja sistem
- Contoh: *off-shift activities, extremely rare down-times, negligible move times, external setups.*

3. Isolasikan waktu aktivitas aktual

10

- Hilangkan waktu tunggu material dan *resources* yang tak ada hubungannya dengan aktivitas
- Contoh: melakukan operasi perakitan seharusnya tidak memasukkan waktu tunggu untuk *item* atau *parts*

4. Cari *common groupings*

11

- Carilah pengelompokan yang “biasa” digunakan saat menghadapi banyaknya variasi dalam simulasi (contoh: ratusan tipe parts, profil pelanggan, dsb)
- Identifikasi kategori umum untuk menempatkan data
- Persentase kasus yang ada dalam setiap kategori dihitung atau diperkirakan

5. Fokus pada esensi, bukan substansi

12

- Abstraksi sistem: sebuah definisi sistem yang digunakan untuk pemodelan sebaiknya mencakup relasi sebab-akibat dan menghindari rincian yang tak berarti (*meaningless*)
- Menggunakan pendekatan “*black box*” untuk definisi sistem
- Perhatikan pengaruh yang ditimbulkan

6. Pisahkan variabel input dari variabel respon

13

- Variabel input menentukan bagaimana sistem bekerja → fokus pada pengumpulan data
- Variabel respon menggambarkan bagaimana sistem bereaksi bila diberikan sekumpulan variabel input
- Variabel respon tidak “mengatur” perilaku model



Langkah-langkah pengumpulan data

14

1. Tentukan kebutuhan data:
 - a. Data Struktural
 - b. Data Operasional
 - c. Data Numerik
2. Identifikasi sumber data
3. Kumpulkan data
4. Buat asumsi bila perlu
5. Analisa data
6. Dokumentasikan dan setujui data

3. Menentukan Kebutuhan Data





Data Struktural (*Structural Data*)

16

- Menyangkut semua obyek dalam sistem
- Semua komponen relevan yang mempengaruhi perilaku sistem
- Termasuk elemen seperti:
 - Entitas (produk, pelanggan)
 - Sumber / *Resources* (operator, mesin)
 - Lokasi (*waiting areas, workstations*)



Data Operasional (*Operational Data*)

17

- Menjelaskan **bagaimana** sistem beroperasi (kapan, di mana, bagaimana kejadian dan aktivitas terjadi)
- Terdiri dari informasi logis atau perilaku tentang sistem (*routings, schedules, downtime behavior, resource allocation*)



Data Numerik (*Numerical Data*)

18

- Menyediakan **informasi kuantitatif** tentang sistem (kapasitas, rata-rata kedatangan, waktu aktivitas, waktu antar kegagalan)
- Untuk sistem baru dimana data tidak tersedia → Sulit ditentukan
- Contoh: kapasitas, waktu kedatangan, waktu aktivitas, waktu antar kegagalan

Penggunaan Kuisisioner dalam Pengumpulan Data

19

Kuisisioner deskripsi sistem:

1. Apa tipe entitas yang diproses dalam sistem?
2. Bagaimana *routing sequence* untuk masing-masing tipe entitas?
3. Di mana, kapan, dan dalam jumlah berapa entitas memasuki sistem?
4. Kebutuhan waktu dan *resource* apakah yang diperlukan untuk setiap operasi dan berpindah?
5. Dalam jumlah berapa entitas diproses dan dipindahkan ? (tentukan untuk tiap lokasi)

Kuisisioner (lanjutan)

20

6. Apa yang memicu perpindahan entitas dari satu lokasi ke lokasi yang lain?
7. Bagaimana lokasi dan *resource* menentukan *job mana* yang dikerjakan berikutnya?
8. Bagaimana *routing* alternatif dan keputusan operasi dibuat?
9. Seberapa sering terjadi interupsi, *resource* apa dan berapa lama waktu yang dibutuhkan bila hal tersebut terjadi?
10. Bagaimana jadwal ketersediaan lokasi dan *resource* ?

4. Identifikasi Sumber Data





Identifikasi Sumber Data

22

- Sistem saat ini: catatan dan orang-orang dengan pengetahuan dari tangan pertama (*firsthand knowledge*) mengenai sistem → sumber yang luas
- Sistem baru: terbatas pada orang yang langsung terlibat dalam perancangan sistem
- Pertimbangkan reliabilitas dan aksesibilitas

Identifikasi Sumber Data (lanjutan)



Sumber data yang baik:

Catatan historis:	<i>Production, sales, scrap rates, equipment reliability</i>
Dokumentasi sistem:	<i>Process plans, facility layouts, work procedures</i>
Observasi pribadi:	<i>Facility walk-through, time studies, work sampling</i>
Wawancara pribadi	<i>Operators (work methods), maintenance personnel (repair procedures), engineers (routings), managers (schedules and forecasts)</i>
Perbandingan dengan sistem serupa	Dalam perusahaan, dalam industri, dengan industri yang lain
Klaim dari Penjual (<i>Vendor claims</i>):	Waktu proses, reliabilitas mesin baru
Estimasi rancangan	<i>Process times, move times (untuk sistem baru)</i>
Literatur penelitian	Penelitian tentang <i>learning curves, predetermined time studies</i>

Identifikasi Sumber Data (lanjutan)



Sumber data yang baik (lanjutan):

Perbandingan dengan sistem serupa	Dalam perusahaan, dalam industri, dengan industri yang lain
Klaim dari Penjual <i>(Vendor claims):</i>	Waktu proses, reliabilitas mesin baru
Estimasi rancangan	<i>Process times, move times</i> (untuk sistem baru)
Literatur penelitian	Penelitian tentang <i>learning curves, predetermined time studies</i>

5. Pengumpulan Data





Pengumpulan Data

26

1. Tentukan aliran entitas (*entity flow*) keseluruhan
2. Kembangkan deskripsi operasi
3. Tentukan rincian insidental dan nilai data



1. Tentukan *entity flow*

27

Tujuan:

- Mendokumentasikan keseluruhan aliran entitas dalam sistem
- Menyediakan bantuan visual untuk mengkomunikasikan aliran entitas pada pihak lain



1. Tentukan *entity flow* (lanjutan)

28

Entity flow vs. process flowchart

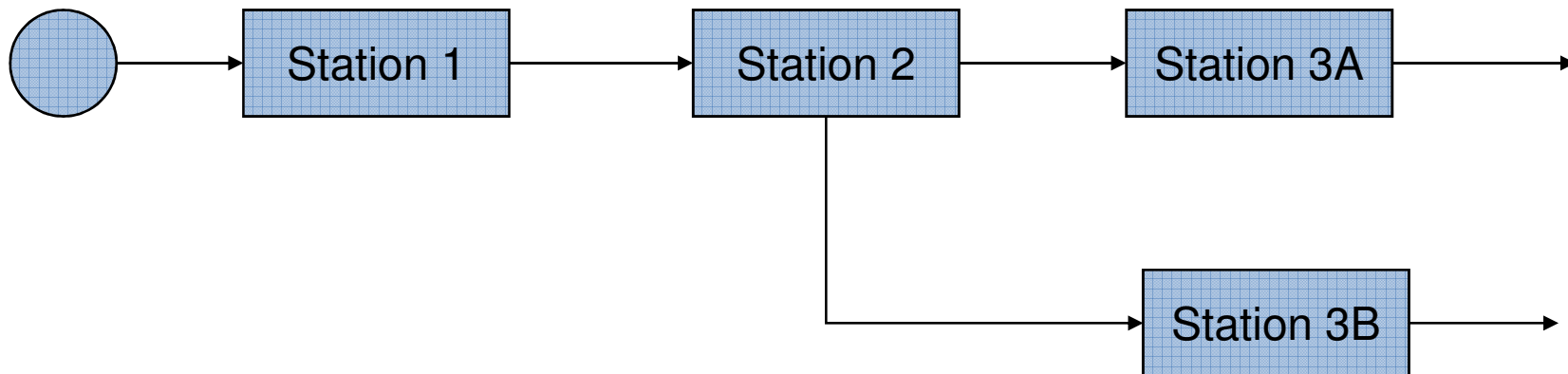
- ***Process flowchart***: memperlihatkan urutan logis aktivitas dimana entitas berjalan dan mendefinisikan **apa** yang terjadi pada entitas, bukan dimana terjadinya
- ***Entity flow diagram***: lebih berupa peta urutan (*routing chart*) yang memperlihatkan gerakan fisik entitas melalui sistem dari satu lokasi ke lokasi yang lain



1. Tentukan *entity flow* (lanjutan)

29

Produk A



Entity flow diagram



2. Mengembangkan deskripsi operasi

30

- Menjelaskan bagaimana entitas diproses melalui sistem
- Dapat ditulis langkah demi langkah (*step-by-step*), naratif singkat, atau dalam bentuk tabel
- Seharusnya mengidentifikasi masing-masing tipe entitas pada tiap lokasi dalam sistem
 - Kebutuhan waktu dan *resource* aktivitas operasi
 - Dimana, kapan dan dalam jumlah berapa entitas diproses selanjutnya
 - Kebutuhan waktu dan *resource* untuk memindahkan ke lokasi berikutnya



2. Mengembangkan deskripsi operasi (lanjutan)

31

Contoh:

Pasien memasuki kantor Dr. Brown dan memberikan tanda-tangan di bagian penerimaan (*check-in counter*) sebelum duduk menunggu panggilan suster.

Pasien diantar menuju salah satu dari tiga ruang pemeriksaan saat salah satu ruangan tersedia. Seorang perawat menemani pasien menuju ruang pemeriksaan, tempat dimana mereka menunggu untuk diperiksa Dr. Brown.

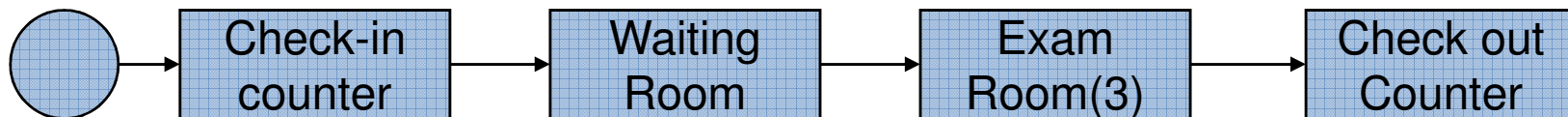
Setelah pemeriksaan, pasien kembali sendiri ke *checkout counter*, dimana mereka membayar dan menjadwalkan kemungkinan kedatangan berikutnya.



2. Mengembangkan deskripsi operasi (lanjutan)

32

Pasien



Contoh: *Entity flow diagram* bagi pasien yang diproses di kantor Dr. Brown



2. Mengembangkan deskripsi operasi (lanjutan)

33

Process Description for Dr. Brown's Office

<i>Location</i>	<i>Activity Time</i>	<i>Activity Resource</i>	<i>Next Location</i>	<i>Move Trigger</i>	<i>Move Time</i>	<i>Move Resource</i>
Check-in counter	$N(1,2)$ min.	Secretary	Waiting room	None	0.2 min.	None
Waiting Room	None	None	Exam room	When room is available	0.8 min.*	Nurse
Exam Room	$N(5,4)$ min.	Doctor	Checkout counter	None	0.2 min.	None
Checkout counter	$N(3,5)$ min.	Secretary	Exit	None	None	None

*Stops to get weighed on the way

Contoh: Deskripsi Proses di Kantor Dr. Brown



3. Tentukan rincian insidental dan nilai data

34

- Untuk mendapatkan model yang lengkap dan akurat
- Untuk sistem saat ini → mengamati data lebih akurat dengan cara mengadakan studi waktu (*time study*) menggunakan sampel

6. Membuat Asumsi





Membuat Asumsi

36

Sebuah model simulasi dapat dijalankan dengan data yang tidak tepat (*incorrect*), tapi tidak dapat dijalankan dengan data yang tidak lengkap (*incomplete*)



Membuat asumsi (lanjutan)

37

- Manfaat simulasi adalah memperkirakan kinerja sistem berdasarkan asumsi yang diberikan
- Kita mungkin menghadapi keadaan dimana tidak ada informasi atau informasi yang ada tidak dapat dipercaya
- Kita harus membuat asumsi untuk kondisi masa depan yang tidak diketahui (*unknown future conditions*)
- Tidak ada salahnya membuat asumsi, sepanjang kepercayaan yang mendasari hasil simulasi tidak melebihi kepercayaan yang mendasari asumsi

Membuat asumsi (lanjutan)

38

- **Analisis sensitivitas (*Sensitivity analysis*):** sejumlah nilai yang diuji pengaruh potensialnya pada kinerja model untuk mengetahui seakurat apa seharusnya sebuah asumsi
- Jika derajat variasi dalam waktu aktivitas khusus memiliki sedikit pengaruh atau tidak memiliki pengaruh sama sekali terhadap kinerja sistem, maka waktu aktivitas konstan bisa digunakan

Membuat asumsi (lanjutan)

39

- Jika tipe distribusi memiliki pengaruh yang dapat diamati pada perilaku model, maka perlu untuk memilih tipe distribusi secara selektif
- Cara mudah melakukan analisis sensitivitas untuk asumsi tertentu adalah menjalankan 3 skenario yang berbeda:
 - Kasus terbaik atau yang paling optimis (*A “best” or most optimistic case*)
 - Kasus terburuk atau yang paling pesimis (*A “worst” or most pessimistic case*)
 - Kasus yang paling mungkin atau tebakan terbaik (*A “most likely” or best-guess case*)

7. Analisis Statistik Data Numerik



Analisis Statistik Data Numerik

41

- Karakteristik data seharusnya dianalisis untuk memastikan kecocokannya untuk digunakan dalam model simulasi:
 - Independensi (keacakan)
 - Homogenitas (data berasal dari distribusi yang sama)
 - Stasionaritas (*stationarity*): distribusi data tidak berubah sejalan dg berubahnya waktu
- Dapat menggunakan program Stat::Fit (terdapat dalam Promodel)

8. Distribution Fitting



Distribution Fitting

43

Data sampel numerik yang telah dikumpulkan dapat digambarkan dalam model simulasi dengan 3 cara:

1. Data dapat digunakan dengan cara yang sama dengan cara saat mereka dicatat
2. Distribusi empiris dapat digunakan
3. **Pilih distribusi teoritis yang paling sesuai dengan data**

1. Distribusi Frekuensi

44

- Distribusi Frekuensi Diskrit
- Distribusi Frekuensi Kontinu

2. Distribusi Teoritis

45

- *Distribution fitting* adalah penentuan distribusi teoritis yang paling sesuai dengan data sampel
- Ada sekitar 12 distribusi statistik yang biasa digunakan dalam simulasi
- Distribusi teoritis dapat didefinisikan sebagai himpunan sederhana dari parameter yang biasanya menentukan dispersi dan densitas
- Contoh:
 - Binomial
 - Uniform
 - Triangular
 - Normal
 - Exponential

3. *Fitting* Distribusi Teoritis terhadap Data

46

- Usaha mengidentifikasi distribusi yang mendasari data yang dibangkitkan
- Prosedur dasar (3 langkah):
 1. Satu atau lebih distribusi dipilih sebagai kandidat untuk menjadi *good fits* bagi data sampel
 2. Estimasi parameter tiap distribusi yang harus dihitung
 3. Uji *Goodness-of-fit* dapat dilakukan untuk memastikan sebagus apa distribusi sesuai dengan data



9. Memilih distribusi bila data tidak tersedia



9. Memilih distribusi bila data tidak tersedia

48

1. *Most likely or mean value*
2. *Minimum and maximum values*
3. *Minimum, most likely, and maximum values*

9. Memilih distribusi bila data tidak tersedia

49

Most likely or mean value

- Contoh dari estimasi nilai tunggal:
 - Sekitar 10 pelanggan per jam
 - Sekitar 20 menit untuk merakit *part* menjadi satu benda
 - Sekitar 5 mesin rusak tiap hari
- Uji sensitivitas sederhana: meningkatkan dan menurunkan nilai hingga 20% untuk melihat perbedaan yang terjadi pada hasil simulasi
- Masalah: tidak ada pertimbangan yang dimiliki untuk variasi yang mungkin

9. Memilih distribusi bila data tidak tersedia

50

Minimum and Maximum values

- Contoh
 - 1.5 sampai 3 menit untuk memeriksa barang
 - 5 sampai 10 pelanggan datang per jam
 - 4 sampai 6 menit untuk menyiapkan sebuah mesin
- Gunakan distribusi normal dengan standar deviasi ditentukan $1/6$ dari jangkauan / *range*
- Masalah: cenderung memperbesar jangkauan lebih besar dari perilaku sistem

9. Memilih distribusi bila data tidak tersedia

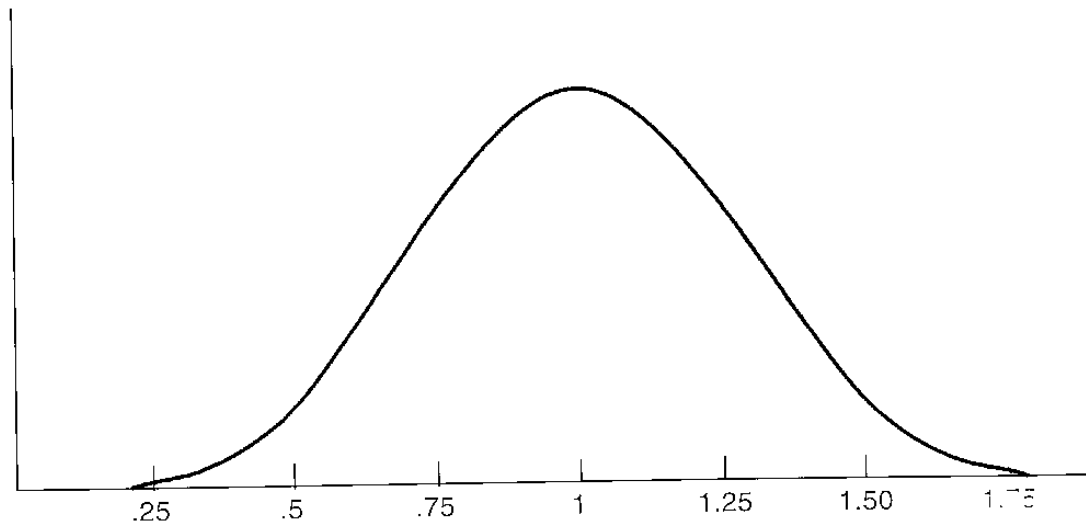
51

Minimum and Maximum values (lanjutan):

- Contoh: untuk *range* waktu .25 sampai 1.75 menit, sebuah distribusi normal dapat digunakan dengan rata-rata 1 menit dan standar deviasi .25 menit

FIGURE 6.21

Normal distribution with mean = 1 and standard deviation = .25.



9. Memilih distribusi bila data tidak tersedia

52

Nilai Minimum, Most Likely, and Maximum

- **Contoh:**
 - Dari 2 sampai 15 dengan kemungkinan terbesar 5 menit untuk memperbaiki sebuah mesin
 - Dari 3 sampai 5 menit dengan kemungkinan terbesar 2.5 menit waktu antar kedatangan pelanggan
 - Dari 1 sampai 3 menit dengan kemungkinan terbesar 1.5 menit untuk melengkapi *purchase order*)
- **Gunakan distribusi triangular atau beta**

9. Memilih distribusi bila data tidak tersedia

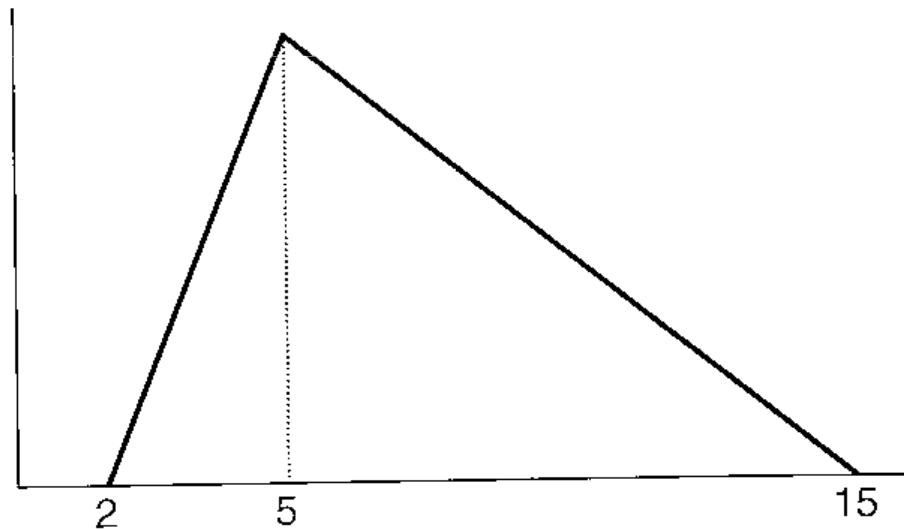
53

Nilai Minimum, Most Likely dan Maximum

- Contoh 2 sampai lima belas dengan kemungkinan terbesar 5 menit untuk memperbaiki sebuah mesin

FIGURE 6.22

A triangular distribution with minimum = 2, mode = 5, and maximum = 15.



10. Distribusi Bounded vs. Boundless



Bounded vs. Boundless Distributions

55

- Distribusi teoritis kontinu dapat dibatasi (rangkap atau tunggal) atau tidak terbatas (*bounded (doubly or singly) or unbounded*)

<i>Distribution</i>	<i>Bounded/boundless</i>
<i>Normal</i>	<i>unbounded (from $-\infty$ to $+\infty$)</i>
<i>Exponential</i>	<i>bonded only on the lower end of the distribution</i>
<i>Uniform and Triangular</i>	<i>Doubly bounded, both lower and upper boundaries</i>
<i>Empirical continuous</i>	<i>Doubly bounded (defined by a finite number of cell intervals)</i>

11. Pemodelan Probabilitas Diskrit Menggunakan Distribusi Kontinu



11. Pemodelan Probabilitas Diskrit Menggunakan Distribusi Kontinu

57

- Distribusi probabilitas diskrit lebih jarang digunakan dalam simulasi dibanding distribusi probabilitas kontinu
- Dalam simulasi, setiap peristiwa untuk membangkitkan nilai dari sebuah distribusi probabilitas diskrit lebih melibatkan pembangkitan nilai integer dibanding nilai nyata
- Salah satu caranya adalah dengan menggunakan distribusi kontinu, dimana nilai pecahan dapat dibuang



12. Dokumentasi dan Persetujuan Data



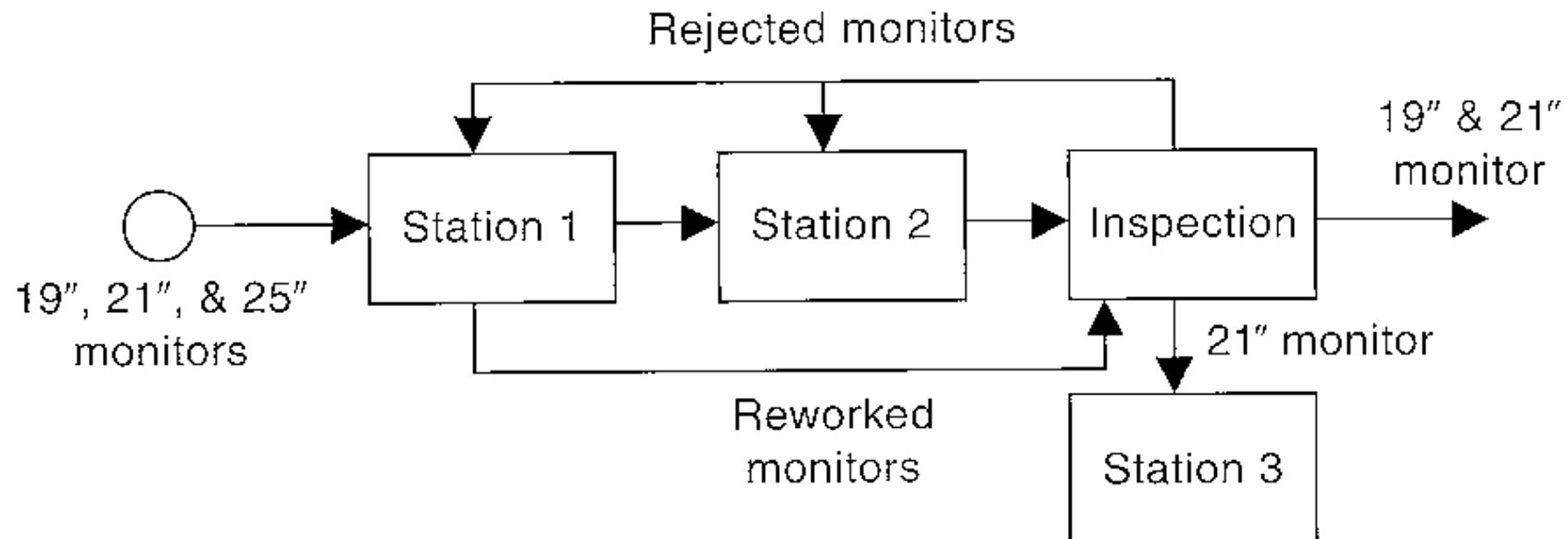
Contoh Dokumentasi Data

59

Tujuan

Tujuan dari studi ini adalah menentukan utilisasi stasiun dan *throughput* dari sistem

Entity flow diagram



Contoh Dokumentasi Data (lanjutan)

60

Entitas

Monitor 19"
Monitor 21"
Monitor 25"

Informasi *Workstation*

<i>Workstation</i>	<i>Buffer Capacity</i>	<i>Defective Rate</i>
Station 1	5	5%
Station 2	8	8%
Station 3	5	0%
Inspection	5	0%

Contoh Dokumentasi Data (lanjutan)

61

Urutan Pemrosesan

<i>Entity</i>	<i>Station</i>	<i>Operating Time in Minutes (min. mode, max)</i>
19" monitor	Station 1	0.8, 1, 1.5
	Station 2	0.9, 1.2, 1.8
	Inspection	1.8, 2.2, 3
21" monitor	Station 1	0.8, 1, 1.5
	Station 2	1.1, 1.3, 1.9
	Inspection	1.8, 2.2, 3
25" monitor	Station 1	0.9, 1.1, 1.6
	Station 2	1.2, 1.4, 2
	Inspection	1.8, 2.3, 3.2
	Station 3	0.5, 0.7, 1

Contoh Dokumentasi Data (lanjutan)

62

Menangani Monitor yang rusak:

1. Monitor rusak dideteksi saat pemeriksaan dan dipindahkan ke stasiun yang membuat masalah
2. Monitor yang menunggu pada stasiun untuk diperbaiki memiliki prioritas lebih tinggi daripada monitor yang baru datang
3. Monitor yang telah diperbaiki dikembalikan ke pemeriksaan
4. Monitor yang telah diperbaiki hanya memiliki kesempatan 2% untuk rusak lagi, dan kalau hal itu terjadi maka akan diambil dari sistem

Contoh Dokumentasi Data (lanjutan)

63

Kedatangan (*Arrival*):

Satu kotak berisi 4 monitor tiba tiap empat jam yang berdistribusi normal dengan standar deviasi 0.2 jam. Probabilitas kedatangan monitor dalam ukuran tertentu adalah:

<i>Monitor Size</i>	<i>Probability</i>
19"	.6
21"	.3
25"	.1

Contoh Dokumentasi Data (lanjutan)

64

Waktu berpindah (*Move times*)

Semua perpindahan (*movement*) terjadi pada *conveyor* dengan waktu sbb:

<i>From</i>	<i>To</i>	<i>Time (seconds)</i>
Station 1	Station 2	12
Station 2	Inspection	15
Inspection	Station 3	12
Inspection	Station 1	20
Inspection	Station 2	14
Station 1	Inspection	18

Contoh Dokumentasi Data (lanjutan)

65

Pemicu Gerakan (*Move Triggers*)

Entitas berpindah dari satu lokasi ke lokasi berikutnya berdasarkan kapasitas yang tersedia pada *buffer input* pada lokasi berikutnya

Jadwal Kerja (*Work Schedule*)

Stasiun dijadwalkan beroperasi delapan jam per hari

Daftar asumsi (*Assumption List*)

- tidak ada *downtimes* (*downtimes* sangat jarang terjadi)
- *Dedicated operator* pada masing-masing workstation selalu tersedia dalam waktu kerja yang terjadwal
- Waktu perbaikan setengah dari waktu operasi normal

Waktu Simulasi dan replikasi (*Simulation Time and Replication*)

Simulasi dilakukan untuk 40 jam (10 jam *warm-up*). Terdapat lima replikasi