

特 輯

統計調查資料의 機械處理

經濟企劃院 統計局

李 圭 高

統計調查資料의 機械處理

目 次

1. 統計局 資料處理의 沿革
2. 業務量
3. 保有 資料處理機械
4. 資料處理機械의 稼動
5. 앞으로의 設置될 EDPS

1. 統計局 資料處理의 沿革

統計調查를 過程面에서 大分한다면 資料의 蒐集過程과 蒐集된 資料의 整理分析過程으로 二分할 수 있다. 이 가운데 蒐集된 資料의 處理問題는 흔히 等閑視되기 쉬우나 資料의 蒐集못지 않게 重要한 것이며 어떻게하면 더 正確하고 迅速하면서도 값싸게 處理할수 있는가 하는 問題는 調查自體의 價値를 左右할 수도 있음은 누구나 認定하는 事實이다.

우리나라에서 統計資料處理에 機械가 利用된 것은 꽤 오래된 일이나 그때는 單純한 手穿孔카드와 分類機에 依存하여 主로 人口센사스(當時는 戶口 調査)結果를 集計하였었다.

完全한 system으로서의 資料處理機械는 1960年 汎世界的으로 實施한 人口센사스와 農業센사스 結果處理를 爲하여 導入된 PCS機械(穿孔카드資料 處理組織)들이 最初의 것이었다.

그 當時 穿孔機 121臺 其他機械 33臺 總154臺의 機械를 總從事員約400名이 晝夜稼動할때는 勿論한 것이었다.

約 4年에 걸친 人口와 農業 센사스의 處理가

끝나기 始作할 時期부터 其他 業務의 機械利用이 始作되었으며 機械臺數도 大幅 늘이고 機構도 漸次 整備되기에 이르렀다.

現在の 資料處理를 爲한 組織은 課單位로서 3個의 係로 나뉘져 있다. 即

1) 課全體의 全般的 計劃과 進度, 作業別 schedule作成과 必要한 programming을 하는 計劃係가 있고

2) 資料의 符號 內容檢査 및 結果表의 整理檢討를 맡은 審査係가 있고

3) 豫定된 計劃과 주어진 作業要領에 따라 直接 機械에서 일하는 集計係로 區分된다.

이 밖에도 便宜上 班이라는 組織을 만들어서 움직이고 있다.

이 組織의 特徵은 어떤 一個의 調査를 完了하기 爲하여 거치는 過程에서 分割한 것이며 따라서 強力한 統制力과 有機的인 調整이 必要한 데 있다.

2. 業務量

業務는 主로 調査統計局 自體에서 調査한 資料를 處理하나 그 밖에도 貿易統計와 같이 財務部에서 周期的으로 보내는 資料를 處理하여 結果를 送付하는 경우가 있고 依頼에 依해서 使用料를 納付받고 處理하는 경우가 있다.

우리가 實地로 處理하는 業務의 種類와 量은 大略 다음第一表와 같다. 勿論 表에 表示된 外에 센사스等 大規模 調査는 每 5年 마다 있다.

(第1表) 주요 통계자료처리 업무량 (1966년)

업 무 분 류	구 분	업 무 량					
		년감처리회수	조 사 표 매 수		카 드 매 수		결과표수
			1 회	년 간	1 회	년 간	
1. 인구주택조사	1회 6월	(125,000매 12일) 1,500,000매	1,500,000매	1,500,000매	1,500,000매	29	
2. 인구주택표본조사	4회 2,5,8,11월		7,000	28,000	32,000	128,000	24
3. 경제활동인구조사	4회 1,4,7,10월		5,000	20,000	20,000	80,000	13
4. 가 격 조 사	4회 1,4,7,10월		1,800부 (200부 3회)	7,200부	60,000	240,000	10
5. 소액물가조사	36회(월3회) 매일5,15,25일		600부	7,200부	81,000	972,000	3
6. 무 동 통 계	12회 매월 초순		12,000	144,000	15,000	180,000	7
7. 범죄 통 계	12회 매월 하순		150,000	1,800,000	150,000	1,800,000	14
8. 의 환 통 계	12회 매월 초순		10,000	120,000	10,000	120,000	5
9. 신용장 통 계	12회 매월 초순		2,000	24,000	2,500	30,000	4
10. 광공업표본조사	12회 매월 초순		3,000	36,000	6,000	72,000	7
11. 농가구매조사	12회 매월 중순		1,500부	18,000부	15,000	180,000	3
계				3,704,000		5,302,000매	

3. 保有 資料處理機械

總臺數 65臺中 自體所有分이 42臺이고 其他 23臺는 賃借分인데 機能으로 봐서 크게 P.C.S와 E.D.P.S로 區分하고 있다. (表2參照)

P.C.S로서는 IBM會社製인 穿孔機(024型),分類機(083型),製表機(101型·419型)와 其他 補助機로 構成되고 있으며 昨年6월에 稼動始作한 E.D.P.S.(電子計算組織)로는 亦是 IBM製의 1401C 6型을 設置하고 있다.

이 電子計算組織은 小型에 屬하는 것이더 可變長語式인 記憶容量 16,000字인 一般事務用 機械이다. 使用하는 programming language는 autocoder를 爲主로 하고 있으며 特殊한 境遇에만 FORTRAN을 쓰고 있다.

附隨機械들은 4臺의 磁氣테이프裝置, 1臺의 카드 리더(Card Read (Punch))와 1臺의 印刷裝置가 있는 簡單한 것이나 機能은 다음 第3表에 表示된 바와 같이 좋은 便이다.

(第2表) DATA Processing Machines Installed at Bureau of Statistics E.P.B. (1968)

Name	Type	Total No.	Owned	Rented
1. Card Punch Machines	IBM 024	27	16	11
"	" 029	2	—	2
2. Card Verifier	" 056	21	12	9
3. Sorter	" 083	4	4	—

4. Collator	" 085	1	1	—
5. Electronic Statistical Machine	" 101	2	2	—
6. Accounting Machine	" 419	1	1	—
7. Reproducing Punch	" 519	2	2	—
8. Summary Punch	" 524	3	3	—
9. Interpreter	" 543	1	1	—
10. Electronic Data Processing system	IBM 1401C6	1	—	1

(第 3 表)

형 격	명 칭	대 수	기 능
1401	중앙처리기구(Central Processing Unit)	1	기록, 판단, 처리, 명령 기억용량 4,000자
4106	부수기억기구(Additional Storage Unit)	1	기억용량 12,000자
1402	카드읽기및천공기구(Card Read Punch)	1	카드에 의한 자료의 해독(800매/분) 결과를 카드에 천공(250매/분)
1403	인쇄기구(printer)	1	결과의 인쇄(600행/분)
729	자기 테이프 (Magnetic Tape Unit)	4	테이프에 의한 자료의 해독 결과를 테이프에 기록(60,000자/초, 75 /초)
	자기 테이프 1권당		길이 2,400. 260,000카드본 수록. 20,000,000 자 수록

4. 資料處理機械의稼動時間

센사스 등의 大規模 業務를 除하면 統計調查資料는 月別 또는 分期別로 處理되고 있다. 1966年度의 例를 보면 各分期의 두째 月의 두째 週는 P.C.S의 Accounting Machine 二交代를 繼續稼動시키고도 時間이 드자라서 三交代를 해야한다는 이야기가 된다. 그래서 統計調查 目標의 規模가 制限되게 마련이며 新規業務나 調查規模의 擴大는 事實上 어려웠다.

그래서 表에 나타나 있는 定規的인 業務以外의 作業에는 여러가지 고충이 따른 일들이 記憶에 남아있다.

1966年10月에 實施한 人口센사스處理用으로 1967年5月에 設置된 IBM1401型 電子計算組織도 充分한 事前計劃의 結果, 設置한 月만 除하고는(始動式等 準備때문에)바로 100% 以上の 稼動率을 보였으며 契約上의 純稼動時間(Net Running Time)을 언제나 훨씬 넘었으며 1967年 末까지 鑛工業센사스, 中小企業實態調查, 資金調查 및 技術系人의 資源調查等 計劃에 없었던 大規模業務만도 5個를 完了하였으리 當初目的

하던 以上の 成果를 얻었다 할 수 있다.

現在는 人口센사스處理를 最初計劃대로 完了하기 爲하여 月平均 純稼動時間 240時間을 움직이며 그거기 爲해서 操作業은 近 300時間을 每月 勤務하고 있다.

第 4 表 통계조사별 평균 및 제표 소요시간 (1966年)

구 분	소요시간(기계)		
	천결공	제 표	계
1. 인구동태조사	16,900시	1,150시	18,050시
2. 인구동태표본조사	600	100	700
3. 경제활동인구조사	1,200	250	1,450
4. 가계조사	4,700	1,150	5,850
5. 소매물가조사	9,000	1,150	10,150
6. 무역공제	2,700	1,100	3,800
7. 벌의공제	11,000	1,450	12,450
8. 외환공제	500	350	850
9. 선용장공제	300	100	400
10. 광공업표본조사	1,000	650	1,650
11. 농가구리조사	1,100	200	1,300
계	49,000	7,650	56,650

(第 5 表) 동지조사별 회계기 소도시간 (1966)

구분	일 수	1, 4, 7, 10				2, 5, 8, 11				3, 6, 9, 12				비 고		
		4	1	2	3	4	1	2	3	4	1	2	3		4	
동계구내업무	1. 인구동태조사									16	10			14	가동비율= 소요시간 가동가용시간 100%	
	2. 가계조사					4	23	23	24	23						
	3. 소매물가조사	5.7		0.6	0.7	5.7		0.6		5.7	0.7		0.6	0.7		5.7
	4. 무역통계		13	26	9			29	10	5		20	19	5		
계	5.7	13	26.6	9.7	9.7	23	52.6	34	33.7	16.7	30	19.6	5.7	18.9		
가동비율 %	16	36	75	28	28	64	147	94	94	47	83	55	16	55		
외부업무	1. 범죄통계			16	16			16	16			16	16			
	2. 광공업표본조사		2				2				2			2		
	3. 의환통계		13				13				13			13		
	4. 소대통계			3				3				3				
	5. 신용장통계	2.3	2.2	2.2	2.3	2.3	2.2	2.2	2.3	2.3	2.2	2.2	2.3	2.3	2.2	
계	2.3	17.2	21.2	18.3	2.3	17.2	21.2	18.3	2.3	17.2	21.2	18.3	2.3	17.2		
가동비율 %	6	48	59	51	6	48	59	51	6	48	59	51	6	48		
총계		8	30.2	47.8	28	12	40.2	73.8	52.3	36	33.9	51.2	37.9	8	36.9	
가동비율 %		22	84	133	78	33	112	205	145	100	94	142	105	22	103	

5. 앞으로 設置될 EDPS

現在 稼動中인 IBM 1401型은 P.C.S中心의 作業에서는 期待할 수 없었던 여러가지 利點이 있었다.

그 主要한 것으로는 첫째 logical operation의 幅이 넓어서 短時間에 data의 누락 또는 重複을 發見하고 內容上의 矛盾된 點을 찾아내서 正確을 期하게 되고 둘째로는 data를 tape에 收錄함으로써 P.C.S에서 일어날 수 있는 여러가지 事故를 防止할 수 있었고 셋째로는 모든 作業, 一例를 들면 分類 計算 製表等이 한 組織內

에서 이루어 지므로 operator의 勞苦를 덜게 되고 error를 낼 可能性이 훨씬 減少된 셈이다.

그러나 業務量의 增加와 制限된 機械의 容量과 性能 또는 經濟性이라는 면에서 볼때 現在의 IBM 1401型의 改善은 設置以前에 벌써 豫見된 일이 있으나 機械의 delivery date문제와 統計局 스스로의 漸進的 姿勢로 于先 1401型이 設置된 것이었다.

그러나 어느 程度의 經驗과 自信을 얻고 今年末에 새로운 型의 EDPS을 設置할 計劃이라 于先 導入된 組織의 規模는 記憶容量 32k biyt의 CPU를 中心으로 한 IBM 360/40型으로 性能은 다음표와 같다.

(第6表) Estimated costs for recommended IBM equipment monthly lease charge.

Machine Type	Model or Feature	Description	Q'ty	Monthly Rental US\$
2040	E	Processing Unit-B2K	1	4,020
	3237	Decimal Arithmetic		126
	4427	Floating Point		105
	7520	Storage Protection		168
	7920	1052 Adapter		235
	6980	1st Selector Channel		395
	4457	1401 Compatibility Feature		560
1052	7	1052 Printer Key Board	1	65
2821	1	Control Unit	1	1,020
	8065	2540 Attach		NC
	8637	Universal Char. Set Adap		15
	1990	Column Binary		105
	5895	Punch Feed Read Control		58
2540	1	Card Read Punch	1	695
	5890	Punch Feed Read		26
1403	2	Printer	1	775
	8640	Univ. Char. Set		10
2803	1	Tape Control Unit	1	750
	7125	7 Trach Compatibility		58
2401	1	Tape Drives	5	1,850
2841	1	Storage Control	1	550
	6118	Record Overflow		10
	4385	File Scan		37
2311	1	Disk Storage Drive	1	605
1316	1	Disk Pack	1	15
				11,036
				11,693